

Where next after Geneva break?

The Geneva arms control talks have adjourned without breaking down, which is something to be grateful for. The negotiators should spend their break deciding where to go from here.

THE prospect that something will come out of the bilateral arms control negotiations at Geneva, far from being dimmed by last week's adjournment, is enhanced because they have not long since broken down. That the negotiations would be difficult has been clear from the start even though those that broke down at the end of 1983 covered some of the same ground. This time, the problems at Geneva have been complicated by its announcement of the US plan to build a defence against strategic missiles, an issue that seems not unduly to have alarmed the Soviet Union during the earlier abortive talks. The best hope now is that the two sides will use the summer to work out the skeleton of an agreement that might be put to Mr Reagan and Mr Gorbachev when they meet, as they intend next November. But even that is a lot to hope for.

The present position is both complicated and delicate. The United States, which spent the years preceding the Anti-Ballistic Missile (ABM) Treaty of 1972 in trying (successfully, in the end) to convince the Soviet Union that ballistic missile defences would be dangerous, is now faced with the task of demonstrating that the opposite is the truth, that a workable shield against hostile missiles would be a boon. Inevitably, the logic must sound thin. Yet the United States is within its rights to insist that research programmes cannot be the subject of international treaties (because verification is impossible), while the Soviet Union can rightly say that some of the testing planned by the Strategic Defense Initiative (SDI) would violate the ABM treaty, as would the emplacement in an Earth orbit of a radar system for detecting hostile missiles, whether or not it was accompanied by a device for shooting down unwanted objects.

Understanding

The signs are that the Soviet Union is now willing to accept that research cannot be prevented by bilateral agreements. What the United States should in return be willing to accept is that successful research cannot be followed by development without violating the 1972 treaty so the need is for some general understanding (not at this stage a treaty) on mechanisms for modifying the ABM treaty by agreement. That, in any case, is about the only way in which the United States could hope to win support in Western Europe for its plans, and also the best way to avoid the increasingly bruising squabbles in Congress about money for SDI.

But how is it possible to agree in advance on procedures for modifying a treaty without knowing what changes will be required? The difficulties are fewer than they appear. The ABM treaty itself is, by its symmetry, a useful guide. In the original version, each side was allowed to protect only two regional targets, one city and one missile field. Each was allowed to build phased array radars, but only on its periphery, and directed outwards. There were restrictions (whose observance is not

easily verified) on the uses that might be made of anti-aircraft missiles. The general principle was that each side was allowed to follow, if it chose, prescribed courses of action. (In the event, the United States decided not to take up the right to deploy anti-missile missiles, the Soviet Union exercised only one of its two options.) So it might sensibly be agreed that any later modifications of the ABM treaty would follow the same principles. The first practical component to emerge from SDI is almost certain to be a recipe for an orbiting early-warning system which, if not accompanied by defences, could do no harm. So why not agree now to agree later on how such a system should be specified, and agree that each side should be entitled to one of them? And so on with the other planned components of SDI, as and if they become practical realities?

The rest of the Geneva agenda similarly cries out for re-examination. The serious flaw in the abortive negotiations of 1983 was that strategic and intermediate-range missiles were rigidly kept separate. The separation is now less complete, which is just as well when it is reckoned that missiles such as the Soviet SS20 and heavy bombers based in Europe have a similar strategic function in a European context. One obvious difficulty now is that so much time has passed since the Salt II agreement (1978) that each side probably needs, for good military reasons, a more modern strategic missile. The US need of something more credible than the MX missile stands out a mile. So here again is an opportunity for reciprocal modification of the existing agreements — permitted modernization coupled, say, with a target number for total warheads reducing steadily over some specific period ahead, something like a decade. There would be opportunities in such an arrangement to use the "build-down" principle in fashion two years ago — one new missile for each two old missiles destroyed. An arrangement like this would let ordinary people sleep more easily in their beds, but also keep the generals happy. But why bother about them? Because an agreement that lets one side feel permanently at a disadvantage will not stick. □

Reaganomics found out

The US economy is heading for stagnation and the need to cut the deficit is more urgent.

BUDGET directors, wherever they work, are not often popular, so that there are rarely floods of tears when they quit their jobs. The departure last week of Mr David Stockman, director of the US Office of Management and Budget since 1981, should be differently regarded. His continued presence at the White House has been for the past three years the best chance that something would be done to reduce the federal government's budget deficit, expected to exceed \$200 million in the year beginning on 1 October. The chance, recently only an off-chance, has no doubt been further diminished by President Ronald Reagan's abdominal operation, while the past week has seen such a confusion of signals about the condition of the US economy from the stock markets and the foreign exchanges that the most resolute members of the US Congress are likely to be given pause. However, this is not the time for masterly inaction.

Biological manuscripts

Contributors are reminded that, with the transfer of the Biological Sciences Editor to the Washington office, it will be helpful if they will in future send **four copies of all manuscripts** offered for publication, either (as at present) to London or Washington.



The past five years have been a topsy-turvy period when it seems to have been possible to make water run uphill. At the outset of the administration, the objective was to cut taxes, in the belief that an industrious industrial economy would make such good use of the extra money that wages and salaries would rise several times as much, thus restoring the federal government's revenues. Taxes were indeed cut, and there has been a remarkable increase in the number of potential taxpayers at work in the United States. But tax incomes have not increased to bridge the revenue gap. So the second and unplanned phase of Reaganomics, most of the past three years, has been a period during which the federal government has bridged the gap of borrowing in the markets, ultimately from overseas.

Inflation has been restrained by the determination to borrow (rather than to print money), economic growth has picked up in spite of historically high interest rates (necessary to tempt in funds from overseas) and the US dollar has been over-valued. The trade deficit, likely to exceed \$140,000 million during this calendar year, is the other side of the same coin, but that is the reason why the pace of economic growth has now declined almost to nothing, at least for the first half of this year. And now, not surprisingly, the value of the dollar has begun to fall. The third phase of Reaganomics has begun. We can only wait and see what it will be like.

The immediate temptation in Washington will be to go slow on attempts to cut the budget deficit. Members of the Congress who have been sharpening their knives will now be given pause by the apparent lack of growth between the first half of 1984 and the first half of this. The Keynesians among them and many others will be saying that a time of stagnation is the wrong time to be taking money out of the system. The hitch in that argument is that if the deficit persists, it will be necessary to increase interest rates again, which may have the same effect, putting a drag on US industry and intensifying the stagnation now in prospect. This is what Mr Paul Volcker, the chairman of the Federal Reserve System, was saying in Washington last week. The alternative would be to embrace inflation, which could easily become the hallmark of phase three.

That is why the United States has no choice but to look overseas for a solution. Japan and, latterly, West Germany have been earning large surpluses in their overseas trade, chiefly from that with the United States, and have obligingly lent that back to the United States. In principle, both governments (and some others) could help avoid the present prospect by allowing their own domestic demand to increase, either by lowering interest rates or (better not) by apeing the first phase of Reaganomics. But none of this is arithmetically possible unless the US budget deficit is cut. How else could Japan and West Germany earn the surpluses with which the US Treasury expects them to bridge its gap? The simple moral is that there should be a negotiation between the three countries principally concerned to decide which deficits should be cut by how much. The pity is that the chance to do this, last month's meeting in Bonn, has been let slip. And next year's meeting in Tokyo is too far away. □

University on brink

The changed University of London could be fun to watch.

THE University of London is, or at least was, a remarkable institution. It is the only really large university in Britain, with more than 20,000 undergraduates. Not so long ago, it had no fewer than a dozen distinct medical schools, and a similar number of medically oriented postgraduate institutions. There is also a handful of specialized centres within the university concerned with matters as different as oriental languages and fine art. Until a quarter of a century ago, the University of London played a kind of authenticating role in British university life, acting as the validating umbrella beneath which sheltered a number of institutions that are now universities in their own right. During the same period, the university also ran single-handed a system of

public examinations for secondary schools throughout what was then the British Commonwealth, a function that has been eroded as much by the wish of Commonwealth countries to run their own affairs as by the university's choice.

The past few years have seen even more rapid change, largely because of the shortage of funds. Five years ago, a committee under Sir Brian (now Lord) Flowers advocated that the number of medical schools should be reduced, by mergers, to a half, and there have indeed been some moves in that direction. More recently, the main teaching establishments of the university have also been reduced in number. Small Bedford College has moved out into the country to merge with another small teaching school, Royal Holloway College, on a site more than twenty miles from central London. Two others (Chelsea College and Queen Elizabeth College) have merged (with effect from 1 August) with King's College to form an institution with more than 6,000 students (plus a medical school). The result is that the university's main teaching functions are for the first time concentrated in a handful of institutions, four large establishments (Imperial, King's, Queen Mary and University Colleges), Birkbeck College which has a special interest in part-time students, the London School of Economics and the amalgamation in the Surrey countryside. (One small college, Westfield, looks like sinking without trace unless it decides in the next few weeks to merge with the new King's College, to be known as KQC for the next five years.)

The objective of these upheavals was originally to save money, and that may eventually be possible. More immediately, however, the crying need is to find capital with which to make the planned amalgamations into effective teaching institutions. With the real funds at the disposal of the University Grants Committee shrinking by between one and two per cent a year, and with little prospect that the British government will dig into its pocket to help out just one university, much of the brave planning of the past few years may come to nothing. The most obvious danger for the colleges (and those who teach in them) is that the grants committee will be forced by sheer penury to shrink student numbers until the existing buildings will accommodate them.

The university is not well placed to defend itself. Administratively it is a mess. Although students at each of the teaching establishments are in theory members of the university as a whole, in practice the large institutions are separate and autonomous. Students are taught where they are recruited, and the different colleges distinguish carefully between each other's degrees. Financially, the colleges are supported by funds from the central federal university (but Imperial College deals separately with the grants committee on the grounds that it is a technological centre of national importance). Everybody agrees that the procedures for sharing out the funds are unsatisfactory, primarily because the customers are never helped to understand the principles on which decisions have been made. The system is also cumbersome; it is a standing joke that the University of London is usually about a year late in replying to the questionnaires that have been a standard part of university administration.

Left to itself, the university would probably collapse under its own weight in the years ahead. Now that the federal university's main teaching functions are no longer distributed among small as well as large institutions, the temptation to circumvent the university machinery, and to deal directly with the grants committee, will be strong. The result would be that the richness of the present diversity within the university would be lost. Most academics within the university know that to be the case, but there is also a limit to their willingness to put up with an administrative system which is not merely cumbersome but which lacks intellectual coherence. The challenge for the university's bosses and in particular for Lord Flowers, the head of Imperial College who takes over as vice-chancellor of the university next month, is that of deciding whether there is a way of keeping the university together and then of persuading those who work in it to follow such a course. It will be a struggle, but one worth making. □

Defence research

Defence to support UK academic research

In a move that could more than double its financial support for university research over the next three years, the British Ministry of Defence (MOD) is embarking on two new collaborative schemes involving universities, industry and its own research establishments. The increase in funds, amounting to anything up to £15 million on top of the £9 million already spent annually on direct collaboration with universities, will come from MOD's research budget, now running at about £380 million per year.

More than one third of MOD's research budget is spent externally, most of it in industry on applied research with specific

University of Geneva

Illmensee's view

WITH the collaboration of others in his laboratory, Dr Karl Illmensee is still hoping to reproduce some of his early experiments on mice before he leaves the University of Geneva in September 1987. The difficulties in doing so, he said this week, are partly the result of continual interruptions caused by the need to attend to administrative affairs connected with the events that led him not to seek an extension of his contract with the University of Geneva (see *Nature* 11 July, p.98).

"The climate here is so hostile that there is no possibility to do research", says Illmensee, and "there is false information" in the internal report that was to have been submitted to the faculty meeting that would have considered extending his contract beyond September 1987.

Apart from interruptions from the university and the press, says Illmensee, the problems of repeating some of the past results are normal scientific difficulties. He has sufficient financial support from the university to attempt to reproduce past experiments and to continue with some of the research that has been interrupted, in close collaboration with others in his laboratory. These include three postdoctoral scientists, one from Hungary and two from West Germany, who have joined the laboratory in the past year.

"My priority", says Illmensee, "is to repeat the experiments of 1982 involving teratocarcinomas." The problem, he claims, is still that of re-establishing the teratocarcinomas. The second priority is to repeat the 1977 production of homozygous diploid mice using a technique that nobody has been able to reproduce. "I have no idea if the time (until September 1987) will be sufficient for me to complete these experiments", he says.

Peter Newmark

applications in mind — not to be confused with the funds spent on development, which in total amount annually to £1,900 million. But a small percentage of the research money goes into strategic ("seed-corn") research. About five per cent of work carried out in the research establishments falls into this category. Externally, MOD supports some 650 research projects among 72 universities, colleges and polytechnics.

For some time, academics suffering a steady squeeze in funds for science have been casting covetous looks at MOD's resources, comparing, for example, its £380 million research budget with the £280 million available to the Science and Engineering Research Council (SERC). Partly in response to concern that it should generally seek to do more for the country's science base, but partly because of its own need to increase the return on its funds, MOD has set up a collaborative scheme for university research grants with several of the research councils, particularly SERC and the Natural Environmental Research Council (NERC).

The new scheme, to start towards the end of this year, should involve only minor change in the councils' grant funding procedures. Applicants will be invited to indicate whether they are interested in, or opposed to, part support by MOD. Following peer review, research staff in MOD will assess appropriate grant proposals and offer up to 50 per cent of the required support. The councils will then decide whether to provide the balance.

Both MOD and the research councils emphasize that the new scheme should not bend council funds towards defence-related research. And an MOD spokesman says that grants with significant defence relevance that fail to obtain council support may still be taken up by MOD. Furthermore, jointly funded research projects will be subject to restrictions on publication only when patents are involved or when researchers move into classified areas. Such occasions "are conceivable", according to the spokesman, "but we will bend over backwards to keep work unclassified. Our aim is not to be bureaucratic." Recent restrictions on US researchers by their Department of Defense, such as the much publicized cancellation of sessions at a recent physics conference, are viewed at MOD with disfavour.

MOD is to circulate every university with a portfolio explaining the scheme and listing the specific areas of primary interest together with names of principal MOD researchers so that interested academics can establish informal contacts. The research councils will issue announce-

ments of opportunity for grant applications, with a deadline of 1 December.

A second scheme for university collaboration is being set up by MOD that will involve the research councils only peripherally through their peer review panels. This scheme has been stimulated by MOD's experience of the Alvey scheme for university/industry collaboration in information technology, and will aim to stimulate such activity in other areas of defence-related research. MOD would support the whole of the university component, to the tune of £1 million in the first year increasing to about £4 million in 1988. By that time, MOD expects to spend £5 million annually on the collaborative grants scheme and £6 million direct to the universities. But if the demand is strong enough, a further £5 million could be available.

Phillip Campbell

Genetic engineering patents

Interferon interference

LEGAL tussles over the commercialization of alpha-interferon have taken a new turn with the filing in Vienna by Biogen NV of a complaint against Boehringer Ingelheim Zentrale GmbH for its marketing of a product containing an alpha-interferon produced by recombinant DNA technology. Last August, Biogen was granted a broad European patent for "recombinant" alpha-interferons. Boehringer Ingelheim is one of several companies that have registered their opposition to the patent with the European Patent Office.

Biogen, of Geneva, Switzerland, and Cambridge, Massachusetts, is suing the West German pharmaceutical company Boehringer Ingelheim over an eyedrop for the treatment of keratitis caused by the herpes simplex virus that has been marketed in Austria since early this year. Acknowledging that the product contains a recombinant alpha-interferon, Dr Dieter Laudien, head of the Boehringer Ingelheim patent department, says that the company, and its subsidiary Bender, does not believe that Biogen's patent can be enforced.

Its opposition, lodged with the European Patent Office in May, is based mainly on evidence that enough information about alpha-interferon and its sequence was known and published before Biogen's patent filing to make Biogen's achievements "non-inventive". By starting its lawsuit against Boehringer Ingelheim, Biogen has clearly signalled that it is optimistic that its European patent will survive the opposition when the matter is considered, probably in a few months time. Meanwhile Biogen is still expecting to receive patent protection in the United States despite the rival patent granted to Hoffman-LaRoche (see *Nature* 21 March, p. 207).

Peter Newmark

Space business

Suppliers ready to break rules

Washington

THE Office of Technology Assessment (OTA) of the US Congress has paid the international space industry the compliment of an economic analysis of its prospects not very different from what might be written about an established industry. In a report published this week, *International Cooperation and Competition in Civilian Space Activities*, OTA says that attempts at international collaboration will probably be blocked by national policies, but recommends that, wherever possible, governments should let space industries develop along unregulated lines.

Much of the report is meant to alert US industry to the emergence of competition in some space activities in other industrialized countries to which, the report says, US policies have not fully adapted. The report identifies the members of the European Space Agency (France in particular) and also Japan as the chief sources of competition during the coming decade.

Space activities are still predominantly within the domain of government, according to OTA. And the big spenders are the United States (which spent \$17,500 million in 1984 on civil and military applications) and the Soviet Union (which is said to have 600,000 people working in the field). US spending is 20 times that of the European Space Agency (ESA), and is the equivalent of 0.47 per cent of gross national product (GNP) compared with 0.08 per cent for France, the next in line.

OTA thinks it unlikely that there will emerge a competitive environment for trade in space-related services and products. "Since France and Japan... have made the decision to join the United States as space powers, it would be wishful thinking to believe that they will fully abide by the trade rules..."

Even within the United States, OTA is not convinced that there will be enough business to tempt private companies into the provision of launching vehicles, as three aerospace companies (Transpace carriers, General Dynamics and Martin Marietta) are in principle equipped to do. The report also says that the foreseeable demand for launching facilities is what it is at present — the need to put communications satellites in orbit — and that other satellite uses, such as for remote sensing and materials processing, "are even more uncertain".

The volume of demand should nevertheless be enough to keep both the US shuttle and Ariane fully occupied. According to figures compiled by Rockwell International and quoted by OTA, the combination of civil and military needs in the United States would be enough to occupy 24 shuttle flights a year by 1988, with possible demand making necessary a further 20 flights a year in the succeeding

five years. These calculations assume that Ariane would be making 10 flights a year from now on.

If these estimates are correct, OTA says, there may be a need for extra launching capacity by 1988. But if the US defence needs are in part met by some other launching mechanism, either the shuttle or Ariane could be under-occupied between then and 1994. Those are the circumstances in which competition might occur.

At what price? The cost of a dedicated shuttle flight has already increased from the \$18 million estimated in 1975 to something like \$80 million now, but OTA points out that this is considerably less than the real cost of the five shuttle flights in 1983, which are estimated to have cost \$375 million each. By 1988, when overhead costs will be spread over more launches, it may still be necessary to charge more than \$100 million for a dedicated launch.

At these rates, OTA says, the shuttle should still be cheaper than Ariane as a launching vehicle for geosynchronous satellites (which is reckoned to amount to \$20,000 for each pound of mass in orbit). But competition may be hindered by the terms on which launching services are offered; the US National Aeronautics and

Space Administration (NASA) is required to ask for all the cost in advance, but Ariane's managers take graduated payments.

OTA offers the US government a series of policy choices on the shuttle. If, the report says, the objective is to capture a large percentage of the commercial demand for satellite launchings, the result will probably be increased demand for use of the shuttle, perhaps even a need for more orbiting vehicles, and increased cost. If the US government were deliberately to encourage private US companies to enter the business, they would probably compete successfully with Ariane, but NASA's role as a space carrier would be reduced. The report also points out that a strategy intended to make the shuttle vehicle to meet "all current and future" needs in the United States would greatly increase the cost of the shuttle programme and probably put Ariane out of business.

On the business of building communications satellites, OTA argues that US suppliers are the cheapest source of hardware (in spite of the strong dollar) and were able to compete successfully for the contracts to build all 20 of the Intelsat consortium satellites to be launched in the period 1984-89. But OTA notes that all 19 European satellites to be launched for European owners during the same period will be built in Europe, as will be the nine planned Japanese satellites to be built in Japan. □

US pharmaceuticals

Monsanto and Searle to merge

Washington

THE planned merger of Monsanto and G.D. Searle, announced last week, will provide a quick shot in the arm for Monsanto's growing interest in drugs for human disease. The move will also give Monsanto access to a distribution network for pharmaceutical products, the development of which is one of the objectives of the recently expanded research facility at the company headquarters at St Louis, Missouri.

Although the Searle family, which at present owns 20 per cent of the stock in G.D. Searle, has made no secret of its wish to diversify its interest since last September, the planned merger appears to have taken the pharmaceutical industry by surprise. The plan is that Monsanto should offer to buy all the outstanding shares in G.D. Searle for a total cost of \$2,700 million.

The combined operation will have total sales of \$8,000 million a year, with Searle rather less than 20 per cent the size of Monsanto. Both businesses have substantial interests very different from their point of overlap in biotechnology, which appears to have been one of the chief reasons for the merger. Monsanto is primarily a manufacturer of plastics and agri-

cultural and general chemicals, while Searle in the past two years has become, at least temporarily, the predominant supplier of the artificial sweeteners based on the di-amino acid aspartame.

Both companies have taken a close interest in biotechnology in the past few years. Searle, which has three major laboratories worldwide, has built up the laboratory at High Wycombe in Britain (with a complement of 600) to be a major centre of research in recombinant-DNA techniques. Monsanto's directly comparable interest is the somewhat smaller unit based at its Life Sciences Center at St Louis, opened formally only last year.

A spokesman for Searle said this week that the two research enterprises were likely only to reinforce each other. With the St Louis laboratory at the beginning of what is likely to be a period of growth and consolidation, the High Wycombe laboratory could well have the advantage of possession of the ground in any reorganization that comes about. But, formally, nothing can be done until Monsanto has made its formal offer for the shares of G.D. Searle, won the acceptance of a majority of the holders and then decided where it will raise the funds with which to pay for them.

John Maddox

Academic innovation

Cambridge science park succeeds

LIKE a well-focused laser beam, the Cambridge Science Park (CSP) is hot in the middle. Its heat is that of fast growth — witness the new construction in every direction, the rising profit margins of the companies and the traffic jams outside the park.

This growth of high technology in a previously "backwater" university town has inspired a flurry of analysis about the reproducibility of the CSP experience and its long-term growth. How fertile is the industry/university link for jobs and increased income? Cambridge has been compared to the early Silicon Valley; does it also need the anchor of big companies with a marketing mentality gauged to the mass market rather than to the specialized niche?

CSP, with its 40-odd companies and 1,000 employees, is only a small part of what is now popularly called the Cambridge phenomenon. Since the early 1960s, some high-technology companies employing 14,000 people have been established in Cambridge, most of them with direct links to the university. CSP's acknowledged role is as symbol, proof and drawing card for what is happening in Cambridge. It is not, however, simply a further extension of that activity. Located just outside the town on land owned by Trinity College, CSP has high-rent buildings, lots of open space, a lake complete with rather antagonistic geese and its own personality.

The scarcity of good industrial sites in Cambridge means that CSP is one of the best places in the area. For this reason it attracts not only Cambridge companies that are growing or moving upmarket but also companies from outside the region which come because of Cambridge's reputation and because Britain's oldest science park is now a good address. They come to get away from "metal-bashing" industrial centres to a town where the air is clean. And increasingly they come to CSP because of its accumulated experience in getting new technology into established industry.

Many CSP companies such as Laser-Scan Laboratories, with its laser-deflection display devices for location of aircraft in distress, use computer technology; the hardware/software companies themselves, however, have so far chosen to stay in town, close to the university computer facilities. Spin-off companies are common among these, as they are in Silicon Valley and on Boston's Route 128. Because of CSP's small size and high specialization, there is less interaction among companies at the park.

An important factor in CSP's growing success is the unconventionality of the "oddball" entrepreneurs. Ideas that might previously never have seen the light of day

outside an academic journal are now not only taking reproducible form but are also making money. At one turn, for example, Dr Felix "Polywater" Franks has turned his expertise to water in supercooling cells. At present, genetic engineers and companies doing drug tests on cell lines freeze their cells to preserve them. "Cells are contrary beasts", says Dr Franks; even if they can be frozen, they are often killed or experience genetic drift. Supercooling, according to Franks, keeps the cells alive and pure.

Another innovator is Andrew Goodfellow of Goodfellow Metals, which specializes in thin films and in changing the surface properties of metals by ion implantation. Goodfellow commercialized one of the first Cambridge inventions to become a success — the Huxley-Goodfellow micromanipulator originally designed by Sir Andrew Huxley. Britain, according to Mr Goodfellow, "breeds a nation of losers". The most important difference at Cambridge is the attitude of mind fostered there.

Dr William Bolton, of Cambridge Robotics, runs a tiny company involved in dedicated electronics for foundries. He sees his role as educating both engineering graduates and foundry management, showing the former that the problems of the assembly line can be challenging and providing the latter with graduates already experienced in technical and financial issues.

The Microelectronics Research Laboratory is so far an anomaly at CSP because it is a university facility. Many of the projects the laboratory works on, according to the director Dr H. Ahmed, are of direct benefit to industry. Examples include research on electron and ion beams, as well as new ways of putting silicon on top of insulators. The laboratory does not do contract research but, instead, collaborates with interested companies, a sharing that often results in joint publications.

Trinity's role as owner of the CSP land places it in a special relationship to the occupants of the park. The juxtaposition of high-technology companies with a college known for scientific excellence might be expected. On the other hand, this marriage of interests is an oxymoron: the universities in Britain have traditionally had nothing to do with applied research, and the timescale of Cambridge, "where things are happening fast if they are promised by next term", is antipathetic to the aggressive speed of high technology.

This timescale has its advantages. Dr J. Bradfield, senior bursar of Trinity College, says that growth was slow before 1980, but because Trinity owned the land already, looked at projects in terms of decades if not centuries, and had no speci-

fic plans for CSP, the park was never in danger of being labelled a failure. In addition, though Trinity definitely wants a close relationship between the CSP industries and the university, it sees no reason to force or formalize them. Here is the crux of the industry/university bridge; according to Bradfield, it is the "accidental and unexpected friends" that provide the finest spark.

Trinity's support of CSP has been an important catalyst for the entire Cambridge phenomenon, lending confidence to the movement according to Dr Nick Segal, whose firm Segal Quince & Partners recently published a report on Cambridge high technology. Dr Bradfield points to the 1960s debate over the role of the university and Cambridge's image as unindustrial (as opposed to Oxford) and quaint. Now, according to Dr Bradfield, CSP has shown that applied research can achieve "results of great interest that also make money".

Acceptance for the changing dynamic of university research and industry is expressed in many different ways. According to Dr Peter Woodsford of Laser-Scan, a scant five years ago his company was not allowed to specify the kind of work a Cavendish research fellow that it funded would pursue. He believes he would have such freedom now.

CSP's growth curve, says Dr Bradfield, is virtually classic in form; a slow first few years followed by a steep rise. CSP insiders agree that CSP, along with the entire Cambridge phenomenon, is still far from reaching a plateau.

As for second-stage growth of the Cambridge phenomenon, one issue is how many small companies will be bought out — a trend already begun between US companies and CAD firms in town. Dr S. Bragg is head of the Wolfson Industrial Unit, whose business is providing bridges between industry and university. He predicts that small enterprises will continue to be started, but cannot say whether existing companies will grow suddenly or stay the same size.

CSP's success will be a function of the model analysts apply. There is no reason why a large number of small companies should not provide at least a certain volume of employment and income, at the same time retaining flexibility by constantly burbling new ideas to the surface. Failing companies would be quickly replaced. It is tempting, however, to apply another model — the leaven of large firms governed by a marketing mentality that thinks in billions, not millions — because this model applies to the older larger science parks in the United States. Large companies offer widespread use of products, lots of jobs and the promise of stability. Already there is an IBM "listening post" on CSP. But for many of the innovators there, big fish moving into their pond would not be a welcome transformation.

Elizabeth Collins

UK laboratory animals

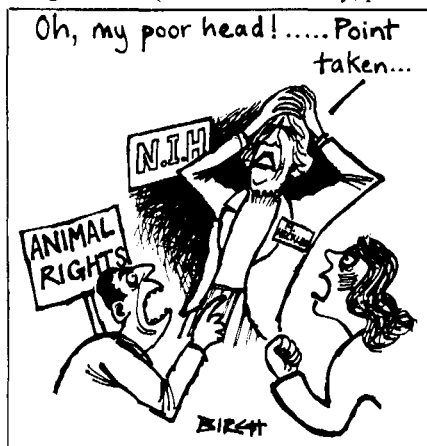
Experimentation reduced

THERE are signs that the pressure of public opinion has caused researchers to think more carefully about experiments on living animals. The number of such experiments in Britain decreased by 4 per cent last year. And although the number of licensed experiments has remained constant, the 3.5 million new experiments reported in 1984 is the smallest since 1959, according to statistics published by the Home Office in its annual report last week (Cmnd 9570, HMSO, £4.70). Just over half of the licences to perform experiments on living animals are held by people in universities and about a quarter by those working for commercial companies; but 53 per cent of the experiments are performed now in the commercial sector and only 22 per cent in universities.

The largest decrease between 1983 and 1984, occurred in experiments on amphibians and reptiles (a 44 per cent fall) and

mice (8 per cent down). Experiments on birds and fish increased but, even here, the total is still less than that in 1980. Even cosmetics testing has been reduced, although only by 3 per cent of the amount in 1983. Of the experiments performed on live animals in Britain, almost all (80 per cent) are either to investigate safety and hazards or to look at body structure and function.

The emphasis on pain placed in the recent white paper (policy document) (Cmnd 9521) on *Scientific procedures on living animals* (see *Nature* 23 May, p. 267)



is continued in the Home Office annual report. The Cruelty to Animals Act requires that animals must be anaesthetized during an entire experiment and, if the animal is likely to be affected post-operatively, killed before recovery. If this frustrates the object of the experiment, the researcher must obtain a special licence from the Home Secretary signed by the president of a learned society and a professor of medical science, as well as reporting each year on details of the experiments performed. The act only allows experiments at all for the purposes of increasing physiological knowledge or for prolonging life and alleviating suffering.

The white paper, which seeks to restrict animal experiments further, is unlikely to be debated in parliament in the next 12 months. This delay has disappointed organizations such as the British Veterinary Association, the Committee for the Reform of Animal Experimentation and the Fund for the Replacement of Animals in Medical Experiments. They feel that the white paper proposals are a necessary first step in the protection of laboratory animals and ensuring that their suffering is minimized. But the government is reluctant to set a timetable for debating such a controversial issue.

Maxine Clarke

• The 21 member states of the Council of Europe have adopted a convention that will lead to better protection of vertebrate animals used for scientific purposes. Bri-

tain will sign the convention "in the early autumn", Home Secretary Mr Leon Brittan says. Britain has played a leading role in drafting the convention and the government intends to use an article "enabling member states to adopt more restrictive measures than those laid down". The convention aims "to balance the needs of research with the consideration owing to live animals; to protect them from pain and suffering; and to ensure that non-sentient alternatives are used wherever practicable". Member states will be required to publish annually statistics on the use of animals. □

Framatome

French company still adrift

FRAMATOME, the French constructor of pressure vessels for the French nuclear reactor network, is still looking for a buyer a year after the collapse of its parent company, the steelmakers Creusot-Loire.

The reason for the delay is simply that Framatome at present does not appear a good buy. Despite the French nuclear success story, most of the reactors that France needs are already built or under construction. The government has now reduced orders to one 1,300-MW reactor every couple of years, and there is unlikely to be an upturn in orders before the oldest reactors, built in the mid-1970s, need replacement in the late 1990s. Exports of French pressurized water reactors (PWRs) have also proved almost impossible. And while Framatome will supply the pressure vessel for the first British PWR at Sizewell, if planning consent follows the public inquiry report due at the end of this year, the French fear that the British market may later be closed. British constructors are expected to get together with Westinghouse to design their own version for the world market, with the result that even if there is an expansion of PWR construction in Britain after Sizewell, Framatome may not benefit. And even if Framatome does get a place, the prospect seems far enough off that the company's immediate problems will not thereby be solved.

Thus the French atomic energy commission, which is now looking after Framatome, would like the company to diversify. Uses for nuclear pressure vessels are, however, relatively few, and Framatome staff are already beginning to fear Framatome's dismantlement. That, however, might prove to be such a political embarrassment for the present government, with nuclear power being rightly seen in France as a national success story, that it can be discounted for the time being. If, however, the right is returned to power at the next elections, it would be a simple stroke to dismantle Framatome — and blame it on the economic failures of the previous government. Robert Walgate

US animal rights

Washington

A FOUR-day sit-in by animal rights protesters at the National Institutes of Health (NIH) ended late last week with the surprise announcement by Secretary of Health and Human Services Margaret Heckler that she was meeting the demonstrators' demand for an end to NIH support of the University of Pennsylvania's Head Injury Laboratory.

The laboratory was the target last year of a break-in by the Animal Liberation Front, during which equipment was vandalized and videotapes of experiments were stolen. Researchers at the laboratory induce brain injuries in monkeys to study how damage and recovery occurs in head-injury accidents.

According to an NIH spokesman, NIH officials were already investigating complaints about violations of its animal care regulations at the laboratory when the sit-in began; the investigators' report was expedited by one day because of the protest, however.

The NIH investigators, who reviewed videotapes and inspected the laboratory, concluded that the researchers had not properly trained or supervised technicians, had performed surgery on animals in unsterile environments and had failed to use proper anaesthesia and analgesia for the animals to minimize pain. NIH left open the possibility that financial support could be reinstated. NIH director James Wyngaarden said he would make a final decision after the University of Pennsylvania has been given a chance to respond. NIH did not criticize *per se* the use of monkeys in the experiments or the experimental protocols.

Stephen Budiansky

French information

New centre for Nancy

NOT before time, France is to reorganize its whole system for managing and distributing scientific information, the French government announced last week, thus responding to a long campaign by the Centre National de la Recherche Scientifique (CNRS), the principal research council, to bring the nation's libraries together in one great documentation system.

According to M. Goéry Delacote, director of scientific and technical information at CNRS, one target will be to catch up with the British Library Lending Division (BLLD), which is now central-

ized at Boston Spa in Yorkshire. Fifteen years ago, said Delacote, the CNRS documentation centre (CDST) and BLLD serviced 300,000 requests each year for photocopies of papers and other items in the sciences. Today, BLLD's computerized systems deliver 2,700,000 a year, while CDST's figure has not changed.

Something had to be done to improve the distribution of papers, a problem exacerbated by the poor financial condition of French university libraries. An opportunity came last year when the government offered to relocate the 350-staff CDST, now in Paris, in the depressed steel town of Nancy. Since many of the staff will probably choose to stay in Paris, this was an offer of jobs to Nancy, and of increased space to CDST.

There will be a new institution, the Agence Nationale de l'Information Scientifique et Technique (ANSTI), whose job will be to coordinate not only CDST but also the many university libraries, the information systems of the medical research council (INSERM), the national collection of periodical catalogues (CCNP) which centralizes the journal lists of 2,300 French libraries, and other sources into one system which should look, as far as possible, simple and unified to the potential user.

According to Delacote, the aim will be not only to increase the volume of papers transported and transmitted in France but also to increase the speed of the various services from their present "rather long" average of 10 days from request to delivery to something more like one or two. "We also want the system to be as comprehensive as possible", said Delacote, so that "95 per cent of the time" a user will find that the paper requested is available.

ANSTI is expected to recommend that the new CDST centre at Nancy will provide the basic document delivery system, will host and develop the French abstracting database Pascal (for the hard sciences) and Francis (for the humanities and social sciences), and provide a centre for information management.

And while France may appear to be coming from behind in this area, it is also moving rapidly into the overtaking lane. Delacote believes a scientific information system like CDST must invest "10, 15 even 20" per cent of its budget in research, and he has put FF25 million (£2 million) of his budget into a project for automatic document delivery dubbed "Transdoc". This experimental electronic archiving and access system is comparing optical digital disks and a microfiche archive run by a computer, and should produce a clear choice of high-volume archiving system "by the end of this year". In this France is running "parallel to or even ahead of US systems".

Robert Walgate

More pork-barrel politics

Washington

DESPITE condemnation by the Association of American Universities, the National Academy of Sciences and the National Science Board, direct appeals to legislators for funding of academic research facilities are continuing, and continuing to pay off. This year, once again, a number of construction grants to universities were approved on the floor of the House of Representatives and the Senate, virtually without debate; the grants had not been submitted to funding agencies previously or subjected to scientific peer review of any kind.

Topping the list this year is Dartmouth College, which, if the House agrees to the Senate's version of the 1985 supplemental appropriations bill, will receive a \$15 million grant for construction projects at its engineering school.

The grant was approved by the Senate without a vote on an amendment offered by Senator Warren Rudman, who represents the state of New Hampshire, where Dartmouth is located. That same amendment also provided \$10 million for a "south wing rehabilitation project" at the Origen Health Science University, located in the home state of Senator Bob Packwood, the chairman of the Senate Appropriations Committee.

Senator Orrin Hatch (Republican, Utah) introduced two amendments to provide support for a \$6 million cancer screening and research centre in St George, Utah, and a centre for the study of health effects of new energy technologies at the University of Utah.

Both projects had been authorized last year in amendments to the Head Start reauthorization bill which Hatch was able to push through in his position as chairman of the Labor and Human Resources Committee. Head Start is a programme that provides special tutoring to disadvantaged children.

Senator Robert Dole (Republican, Kansas), the Senate's majority leader, secured \$200,000 for a remote-sensing research project at the University of Kansas. The one contribution from the Democratic minority in the Senate was a \$500,000 amendment proposed on behalf of Senator Edward Kennedy of Massachusetts for a Chinese-US student exchange programme at Tufts University in Boston.

Earlier this year, Senator Alfonse D'Amato (Republican, New York) pushed through a measure to set aside \$1 million for a computer research programme at the University of Syracuse; the money comes out of a \$25 million university research initiative at the Department of Defense that was to be used entirely for investigator-initiated, peer-reviewed grants for basic research.

Stephen Budiansky

West Germany

Support for biotechnology

Hamburg

BY providing DM 1,000 million between 1985 and 1989, the Minister for Research and Technology, Heinz Riesenhuber (CDU), hopes to enable West Germany's biotechnological research to perform "jumps of perception" as chemical research did 150 years ago.

One principal task is to expand the Gesellschaft für Biotechnologische Forschung (GBF) into a national centre of biotechnology. Another is to promote the gene-centres in Cologne (plant cultivation), Heidelberg (immune defence, antibodies for cancer and rheumatism diagnosis) and Munich (chemical synthesis of nucleic acids) and to support two new centres for biotechnological process engineering.

Apart from financing basic research, Dr Riesenhuber, generally known as an obliging supporter of industrial research, plans to spend DM 120 million on a programme to further the biotechnological industry (*Programm zu Förderung der biotechnologischen Industrie*). Up to 40 per cent of expenses for projects will be refundable by the ministry up to a maximum of DM 600,000. This programme covers developments in cell and tissue culture techniques, genetic engineering, biotechnical procedures with human, animal and plant cells or manipulated microorganisms, new biological systems of plant protection and the use of enzymes in medical and food research.

Development of new apparatus and devices for biotechnological production will also be promoted in the new programme. Dr Riesenhuber is optimistic about the prospects.

The total value of biotechnological products in West Germany came to DM 17 million in 1983, compared with an estimated DM 40,000 million worldwide.

Jürgen Neffe

Bhopal

Birth defect fears allayed

New Delhi

ALTHOUGH there was an increase in the number of abortions, the gas leak tragedy in Bhopal last December did not result in the birth of handicapped babies as had been feared. Some 3,000 women were thought to be pregnant when methyl isocyanate (MIC) leaked from the Union Carbide pesticide factory on 2 December 1984, killing over 2,000 people. The Indian Council of Medical Research (ICMR), which has been following up 1,500 cases of pregnancy, says that by June, 500 babies had been born. Of these, five were born dead and three had congenital defects, an incidence statistically no higher than in a normal population. Most babies were low in birth weight (2 kg) but this too is not uncommon in India.

The rate of 17 abortions for 100 pregnancies was, however, above the national average. Dr V. Ramalingaswami, head of ICMR, says that this was not unexpected and that abortions increase by as much as four times during famine or epidemics, and the Bhopal tragedy was a comparable situation.

Soon after the accident, the state government of Madhya Pradesh considered compulsory termination of all pregnancies in the affected population. The proposal was dropped on the advice of ICMR which had no scientific data to warrant this step.

ICMR will, however, watch for any abnormality in the 1,000 babies expected to be born in the next two months from mothers who were in their first trimester of pregnancy at the time of the accident. "If MIC was really toxic to the fetus, it will be this group that needs watching", Dr Ramalingaswami said. According to his deputy, Dr S. Sriramachari, scattered anomalies in different organs would be a sure sign of MIC effect. But on the evidence so far, ICMR does not expect to find any anomalous births. In women yet to be delivered, ultrasonic scanning has shown fetal growth retardation in 15 per cent of cases, not unusual in India.

Meanwhile, ICMR has completed a health survey of 90,000 people, half the population that was exposed to MIC. Most are still undergoing treatments for respiratory and asthma-like problems, weakness of limbs and inability to carry on their usual work. Patients who had apparently recovered are said to be returning to clinics for recurring ailments. Some 1,200 patients are receiving injections of sodium thiosulphate (STS) which, according to Sriramachari, has been effective. STS is an antidote to cyanide poisoning but no one is sure why it works in MIC victims. It is possible that some MIC broke up into hydrogen cyanide when it escaped from the factory or that it produced cyanide radicals inside the body, after being inhaled. In any case,

ICMR's follow-up of STS-treated cases had revealed that STS mopped up the cyanide from the body and excreted it through the urine. ICMR is not sure how long to keep the patients on STS as they are found to develop symptoms once the injections are withdrawn.

While there has been no breakthrough in treatment, researchers claim to have obtained new information on the mechan-

ism of MIC poisoning. According to Sriramachari, analysis of blood samples from 60 cases has clearly shown that MIC attaches "to the terminal amino group of the alpha chain of the globulin part of the haemoglobin by a process of carbamylation".

This mechanism, he says, "explains the observed clinical features such as respiratory distress, associated with cherry-red venous blood, and toxicity due to failure of utilization and removal of CO₂ and consequently, deoxygenation".

K.S. Jayaraman

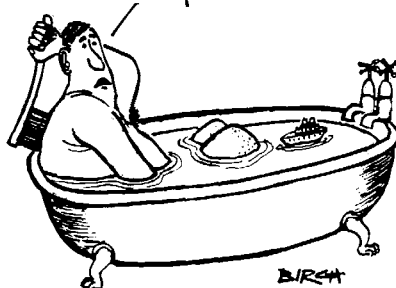
Eureka

Party launches agenda only

EUREKA, the French-led programme to unite European high-technology industry behind a handful of new products from driverless tractors to supercomputers, came into official existence on Wednesday last week. But exactly what Eureka is remains obscure, a situation that led the Paris newspaper *Le Monde* to comment that the decision was like film-maker Woody Allen's remark, "the answer is yes, but exactly what was the question?"

Nevertheless, officials present at the Paris meeting of ministers from 16 European states and the European commissioner for research count the broad agreement on Eureka a success. They point out that in the afternoon and evening allotted

We'll think of what it is soon...



to the discussion, there was hardly time for each delegation to make its own position statement, let alone to reach agreement on "points of detail".

Among the details not yet settled are the fundamental issues of who pays, how the programme is to be managed and on exactly what criteria projects are to be selected. The French President, M. François Mitterrand, put FF1,000 million (£82 million) on the table for 1986. West Germany made "ambiguous" noises about finding DM300 million (£75 million). But otherwise, there was silence from the 17 delegations.

There was, however, a frosty joke between Jacques Delors, president of the European Commission, and Roland Dumas, the French foreign minister. Delors, who has been put out by French insistence that Eureka should not fall into the hands of the "bureaucratic" Commission, has an ambitious technology programme of his own, and said he would

dearly like 20 per cent of national research and development budgets, to run joint European projects. Dumas replied that he would dearly like 20 per cent of the Commission's budget, to support Eureka. Stalemate.

Predictably, Britain was adamant that no special funding was required. Foreign Minister Geoffrey Howe claimed that Europe spends more on research and development than Japan, but gets less out of it. Therefore the problem was not one of research budgets but rather of opening markets. He therefore suggested, as part of a British package of projects for Eureka, that any products created through Eureka collaborations should receive what he described as "Eurotype" certificates — which should guarantee the product free passage through the many restrictive trade barriers that exist in Europe. This proposal is said to have been well received.

For now, the management structure of Eureka remains an open question, despite a strong play by France for a "small" office to be set up in Paris. Dumas said after the meeting that "some countries" had asked France to host the Eureka secretariat, but for the moment Eureka remains without a home.

In the brief final communiqué from Paris, the European Commission is described as acting "in strict consultation" with the Eureka action group. The communiqué also says that Delors' proposals "adopted" by the recent summit meeting in Milan will be "put into action", but what exactly this means will become clear only in future European Council meetings, but it must be doubted whether it will go anywhere near as far as the doubling of Brussels research budgets that has long been in the Commission's mind.

Otherwise, the communiqué says that "particular importance" will now be attached to "encouraging and stimulating the formulation of concrete projects among the industries and research centres of Europe", and to "conceiving the appropriate financial mechanisms". That is what everyone, industry included, is waiting for.

Robert Walgate

Poland

Environment at long-term risk

POLAND will need at least two decades of "huge social efforts and financial outlay" to restore an ecological balance that will satisfy the nation, according to Stefan Jarzebski, head of the Office of Environmental Protection and Water Economy. Presenting a report on the state of the environment over the past five years, Jarzebski told the Sejm (Parliament) last month that the losses caused by the degradation of the environment are now considerably greater than what is spent on attempting to counteract them. Decades of neglect of the environment have produced "calculable economic losses" — smaller crops, smaller yields of forest products, higher costs of water treatment for industry and domestic use.

A state inspectorate of environmental protection established in January 1980 has been insufficient, as the Sejm debate and resolution on the report made clear, to halt the deterioration of the environment. Since the law came into force, "extraordinary threats" to the environment have almost quadrupled, from 396 incidents in 1981 to 1,395 in 1984. Factories find it cheaper to pay fines for emitting pollutants than to take steps to eliminate the sources of pollution. (In 1984, almost 3,000 enterprises were fined a total of more than 2,500 million zloty.) Many important anti-pollution measures recommended under the 1980 law could not be put into practice, or even started, the Sejm noted, presumably as a result of the economic crisis. It should be noted, incidentally, that the closing of the Skawina aluminium refinery near Kraków in January 1981, an achievement to which Jarzebski attached great importance in his report, was not a government initiative at all. Demands to close the plant (which was discharging hydrogen fluoride into the atmosphere) came from the Kraków Ecological Club, one of the many private initiatives which sprang up in the wake of Solidarity, and which, with the backing of the Archbishop of Kraków, Cardinal Franciszek Macharski, managed to convince the relevant authorities that the cost of keeping the plant open, in terms of damage to the local environment and to the health of the population, was greater than the loss to the economy produced by closing it.

In spite of the closure of the Skawina plant, the Kraków area still remains one of Poland's black spots, the others being the Upper Silesian industrial belt (centred on Katowice), the Legnica-Głogów copper belt and the Gdańsk and Puck gulfs. Here, the Sejm decided, the deterioration may be classed as "ecological calamity". Twenty-three areas were designated as being at ecological risk. Even in the capital, Warsaw, householders have to boil their drinking water, and in spite of this

precautionary measure, digestive disorders, urinary tract illness and cancer are on the increase.

The law of January 1980 was one of the few positive responses to environmental problems of the Gierk regime which tended to deal with such matters by censoring public discussion of them. The present government seems at least prepared to discuss them, perhaps because it has a ready-made excuse in Poland's continuing economic problems for any failure of delay in taking action. The Sejm's response to Jarzebski's data was to call for a national programme on the protection of the environment up to the year 2010, a plan which is expected to take at least two years to complete. In the meantime, "undertakings serving protection of the environ-

ment", says the Sejm, should be written into the 1986-90 Five Year Plan.

How far these recommendations will be accepted and implemented is, of course, a matter for the government and the industrial managers.

● Three leading Polish dissidents, Jacek Kuron, Adam Michnik and Tadeusz Jędrzak suggested, last February, that independent Committees for the Protection of the Environment of Silesia should be established, the underground newsletter *Przegląd Wiadomości Agencyjnych* reported on 2 June. The report drew special attention to the high child death rate from cancer in Silesia due to the pollution of the environment, but noted that the majority of the population seemed unwilling to face the reality of the threat. No concrete action seems to have been taken on the proposal, and Michnik is now serving a prison sentence and Jędrzak awaiting trial on political charges.

Vera Rich

Soil Survey

Budget squeeze hurts projects

THE dwindling staff of the Soil Survey of England and Wales is still in the dark about its future. The unit, based at the Rothamsted Agricultural Station, will nevertheless have to decide by the end of October how many members of its staff will be retained beyond the beginning of the next financial year on 1 April 1986, when its budget will be cut in half.

The problems at the Soil Survey appear to be an almost classic illustration of how the Rothschild principle for research support can turn sour. The survey, formally established in 1939, and now constitutionally a part of the British Agricultural and Food Research Council (AFRC), has been supported since 1972 by research commissioned from the Ministry of Agriculture, Fisheries and Food, most recently (in 1983-84) to the tune of £1.8 million a year. In January this year, however, the ministry let it be known that the annual grant would be cut in half (to £850,000) in the year beginning next April. Nothing has been proposed for the succeeding year.

One inevitable consequence of this uncertainty is that the survey is losing staff. By the beginning of this year, largely because of declining research commissions, the staff at the survey had shrunk to 70 (from a peak of 85 in 1976); now it has fallen to 64, largely by allowing vacancies to go unfilled.

The decision about next year's budget was announced by Lord Belstead, an agriculture minister, in the House of Lords on 29 January, who explained that the objective of the reduced grant from next year was the British government's belief that "this is an area which ought to be commissioned by those who use the work". Much in that spirit, AFRC has commissioned a consultants' study of the opportunities for

persuading the Soil Survey's users to pay for the work it does. The report of that study, not yet formally presented, is understood to take the view that the survey could become self-supporting only on a smaller scale, but that a full 50 per cent reduction would not be needed.

The Soil Survey appears thus to be one of the first casualties of the work of the Agricultural Priorities Board, appointed a year ago to replace the previous Joint Consultative Organisation for agricultural research. Lord Belstead said last January that the decision to cut the budget of the Soil Survey had been taken after consultation with the board.

In the past few months, some outside contract work has come the way of the Soil Survey, which says it has been able to raise £175,000 to support certain aspects of its work. "That's six people's jobs for a year", one staff member said. Meanwhile, the pattern of the survey's work has been modified so as to concentrate on projects that could be rounded off by next year. For the time being, work on the 1:50,000 soil map of England and Wales, a project planned to be completed only in the 1990s, has been put on the back burner.

In principle, according to ministry sources, there is no reason why AFRC should not shoulder some of the burden of keeping the Soil Survey alive, although the council itself is under serious pressure from a reduction of its budget. The most likely help from the survey is some adjustment of the cost of premises at Rothamsted, now valued at some £350,000 a year. Paradoxically, nothing has yet been said about the comparable soil survey operation in Scotland, based at the Macaulay Institute, which was reduced by a fifth in 1982 but whose more distant fate will not be decided until the end of August. □

Pain and laboratory animals

SIR — The government's proposals to update the Cruelty to Animals Act 1876 (*Nature* 23 May, p.267) have been widely reported as tightening up the present deplorable situation, a view we strongly challenge.

On the question of pain, the government has proposed that the "severity" of an experiment should be balanced against its potential benefit. Three categories of severity are envisaged: mild, moderate and substantial. In addition, the government has set a "top severity limit" where "an animal which is in severe pain or distress which cannot be alleviated shall be immediately and painlessly killed".

Since it is impossible to devise a method of measuring pain in animals it follows that what may be considered "severe" pain or distress by one person may be dismissed as mild, moderate or substantial by another. Some government departments and officials share our view on this matter. For instance, the Home Office Departmental Committee of Inquiry on Experiments on Animals 1965 (the Littlewood Committee) concluded that "it is not as a rule possible to assess degrees of real pain in animals" (para. 187) and "it is often much more difficult to detect physical pain in an animal than in a human patient" (para. 174).

The subjective assessment of "severe" pain or distress renders such pain conditions meaningless to the animals, while at the same time allowing the government

and experimenters to pretend that there are permissible levels of pain, and that animals are thereby protected against excessive suffering. Furthermore, how is the importance of the experiment to be defined and does this refer to therapeutic, commercial or scientific considerations? Take for example the LD50 test in which animals are deliberately poisoned to death, which is clearly one of the most painful animal experiments. On the basis of the government's proposals, LD50 would have to fulfil a really important need, since the test causes so much suffering. But LD50 is generally recognized as bad science, so would the government refuse to allow the test to be performed? Or would the "severity" of LD50s be redefined as "moderate" or "mild" to overcome these difficulties? The latter course seems highly probable since the government has stated that it will not prohibit the LD50 test.

But the *idea* of prohibiting experiments, such as the use of animals to gain manual dexterity, is already established in the Cruelty to Animals Act 1876 and could readily be extended to other types of experiments, as a positive step towards a completely ethical system of research and health care.

ROBERT SHARPE

*Mobilisation for
Laboratory Animals,
51 Harley Street,
London W1N 1DD, UK*

Romanian reactors

SIR—We would like to provide some additional information on the Romanian nuclear programme (*Nature* 2 May, p.4). There is nothing misleading, as you put it, about our references to two plants now under construction in Romania.

The reactor civil works are well advanced on the first three units, the main excavation for the fourth unit has been started and site clearing has been commenced for the fifth. The office building and some of the permanent service buildings have already been completed and partly occupied.

The cooling water intake pumping house is well advanced and the turbine foundations are reaching completion for the first and second units. The pipe fabrication shop, welder training and electrical shops are already in position.

In terms of manufacture, all the major orders have been placed. Babcock and Wilcox has already shipped heat exchangers and the first calandria or reactor vessel, the largest single component, is scheduled to be shipped from Canada before the end of this year. The arrangements between Atomic Energy of Canada Ltd (AECL) and the Romanians are that AECL will provide licence information so

that the Romanians can undertake a significant amount of the work themselves on the subsequent units.

The drawings and specifications have been delivered to the Romanians and they have a large engineering and drafting team of around 1,000 key staff who are now fully familiar with the engineering details of the CANDU reactor.

The fuel manufacturing shops in Romania have been completed and fuel is being manufactured and stored.

Commitment to the programme is evident to the highest levels in Romania, and this was obvious at the recent visit of President Nicolae Ceausescu to Canada. It included a tour of the Gentilly CANDU Nuclear Power Plant during which Mr Ceausescu entered the containment building and inspected some of the components while the reactor was at full power. He also reviewed the Romanian plans to commit further units and stated that the site for the next five reactors would be chosen before the end of this year.

A. I. SMITH

*Project Director, Cernavoda
Atomic Energy of Canada Limited,
CANDU Operations,
Sheridan Park Research Community,
Mississauga, Ontario L5K 1B1,
Canada*

The nose knows

SIR — Once again Gaylarde has taken it upon himself to "summarize briefly" my hypothesis (*Nature* 315, 92; 1985), and once again I must urge interested readers to read the original instead (*Nature* 311, 515; 1984). It is true that "the inhabitants of the dry desert regions of North Africa and the Middle East, where natural conditions provide the most viscous air to be found on Earth, are not noted for short noses with wide nostrils", but this is not evidence of a lack of effect from climate. People in hot dry places have had to adapt to dust, because hot dry places produce dust more easily and because hot dry air supports more dust. More properly, dry particles float in dry air while humid particles (humidity gained from humid air) sink. That is why people in dusty climates have dust-trapping noses.

Regarding the irrelevance of Gaylarde's advertised "familial virtue", I only wish to be clear that I do not claim familial virtue, or even personal virtue. I claim to have made an argument which is logical and empirically reasonable. Gaylard has decided that I am either a fascist or a would-be fascist (*Nature* 313, 425; 1985 and see also *Nature* 314, 398; 1985). I would prefer to hear his thoughts about my argument, but it seems that he will never calm down enough to get it straight.

JOHN HARTUNG

*Department of Anesthesiology,
Downstate Medical Center,
State University of New York,
450 Clarkson Avenue,
Brooklyn, NY 11203, USA*

Safety cabinets

SIR — In commenting on the relevant advantages of Class I and Class II safety cabinets (20 June, p.626), R. P. Clark deals only with the containment of aerosols. Much of my 40-year professional career as a medical microbiologist has involved work with high-risk pathogens including smallpox, lassa fever and simian B virus, and I have used all three classes of cabinets for this work. My experience leads me to say that for any pathogen in which the risk to the operator arises mainly from needle-stick type of injuries, Class II cabinets afford a significantly higher level of protection than Class I simply because of the better visibility which they afford and the better ergonomic design which is possible for Class II cabinets.

In my view it would be irresponsible to work in a Class I cabinet if a Class II cabinet was available, for this type of work.

K. MCCARTHY

*University of Liverpool,
Department of Medical Microbiology,
Duncan Building,
Royal Liverpool Hospital,
Prescot Street, PO Box 147,
Liverpool L69 3BX, UK*

Why celebrate laser birthday?

Lasers were first constructed 25 years ago. An historical account and some recent work appear in this issue. There can hardly ever have been an instrument that so quickly became pervasive.

WHY celebrate the 25th anniversary of the invention of a scientific instrument, the laser? Why not instead be patient, and wait for the more conventional half-century? Or, since lasers are mostly tools rather than objects of interest in their own right, why not celebrate the discoveries their existence has made possible rather than the means by which these ends have been successfully attained? These questions are all properly provoked by the group of articles on pages 307 to 330 of this issue. What follows is the explanation.

First, the supposition that lasers are merely tools for doing an admittedly huge diversity of jobs is easily made but is also false. With all that has happened in the past 25 years, it tends to be forgotten that the development of the first working laser was also the first occasion on which it became possible to study coherent electromagnetic radiation at or near the wavelength of visible light. By contrast, the study of coherent microwave radiation was in no sense dependent on the earlier invention of the microwave analogue of the laser, called the maser.

Both devices were also, as it happens, in their time the most vivid practical demonstrations of the reality of the phenomenon, originally inferred by Einstein during the development of his quantum theory of radiation (in 1912), that an excited atom capable of emitting a photon of a certain frequency will be stimulated to do so by radiation of that frequency or something like it, and that the likelihood of stimulated emission will be proportional to the intensity of the stimulating radiation. Although there was never any reason to quibble about Einstein's conclusion, which was enthusiastically used in the 1920s by R.H. Fowler and A.S. Eddington (for example) for the calculation of the transmission of radiation through stellar atmospheres, and although the direct calculation of the Einstein coefficient became one of the first successes of the wave mechanics of 1926, it always helps to turn a phenomenon into a fact of life that there should be a direct demonstration of it.

The device constructors now have an open licence. Their objective seems simply to make every possible kind of electronic excitation into the working principle of a laser, whence the emergence of solid-state semiconductor lasers (where the excitation is that of electrons into the conduction band). The outcome is coherent radiation because the stimulated emission

is always in phase with that which stimulates it. The practical difficulties are always those of producing a sufficient supply of excited entities, electrons in atoms or in semiconductors, to swamp incoherent spontaneous emission of the same radiation. So pumping (from the ground state to the excited state) is where the designers spend their energy.

Weird and wonderful things are done in the process, as shown by an account in the current issue of *Applied Physics Letters* of an attempt by R.S. Taylor and K.E. Leopold of the National Research Council of Canada to extend the pulse-length of a xenon/chlorine excimer laser (*Appl. Phys. Lett.* 47, 81; 1985). Here the excited state is that of an ionized atom, the lower level is a loosely-bound outer electronic state and the energy difference corresponds to ultraviolet radiation. Pumping such a device entails the creation of free electrons, which means an electric discharge.

But how can such a phenomenon be sustained for a reasonable length of time, say a microsecond? The answer offered is a double bank of capacitors discharging through one half of a rudimentary transformer, one quickly (in a few tens of nanoseconds) at high voltage and the other more slowly, through the laser gas made electrically conducting by the trigger pulse. The engineering of this equipment entails arguments of the kind now most commonly encountered among those who organize the discharge of the energy of capacitors into the plasma contents of thermonuclear fusion machines (but the laser is on a much smaller scale).

Taylor and Leopold claim an efficiency of 2 per cent for their device, with a 0.4 microsecond pulse of ultraviolet radiation, and indeed look forward to the continuous operation of devices like theirs. The US Strategic Defense Initiative will be interested. Taylor and Leopold calculate that they have extracted 1.5 J of power from each litre of lasing gas in a pulse lasting only 0.4 microseconds, which corresponds to 3.75 MW at continuous operation. Pumping the materials from which X-ray lasers may be made has still some way to go.

These are the practical problems; what are the benefits? Even after so short a time as 25 years, it is easy to overlook, or to take for granted, the way in which the use of lasers has transformed whole areas of investigation. Chemists concerned with the mechanisms of chemical reactions

were among the first to appreciate the value of devices that could be used to excite molecules in specific ways and in times short compared with those of chemical transformations and interactions. One tangible outcome, for example, has been a knowledge of the role of molecules of chlorophyll in photosynthesis that was previously entirely inaccessible. Similarly, there is now a whole new field of spectroscopy, that of the highly excited atoms called Rydberg atoms, whose existence has been made possible only by the availability of tools that can be used in the preparation of the Rydberg states. The value of being able to make measurements of the properties of the species involved in, say, the upper atmosphere or in stellar atmospheres should be huge.

The most casual readers of the literature must now be aware of the pervasive influence of lasers in most fields of research. (The issue of *Applied Physics Letters* carrying the article by Taylor and Leopold has no fewer than eight other accounts of schemes for making novel lasers function, but the journal is the device-constructor's running commentary.) Recent demonstrations that lasers, by means of the momentum transferred in the absorption of photons, can be used to rob atoms of virtually all their translational velocity ("cooling" is the term) points the way to a further increase in the accuracy of atomic spectroscopy and thus to the definition of frequency standards, or time standards, more accurate than even those now used (see B.W. Petley, *News and Views* 27 June, p. 716).

It is also now a matter of common knowledge that the stability of certain geological structures, critical parts of the San Andreas fault, for example, are monitored by geodetic lasers, that other instruments are used for the direct measurement from the ground of the constituents of the upper atmosphere (LIDAR) and that the accumulation of data from Earth satellites and from the corner reflectors left on the Moon during the Apollo missions will eventually yield a more precise test of current understanding of the dynamics of the Solar System, relativistic corrections included.

So there is plenty to celebrate. That lasers would turn out to have such an influence in so many different fields was foreseen 25 years ago. The surprise is that so much has happened so quickly.

John Maddox

Muscle contraction

Weak and strong crossbridges

from Malcolm Irving

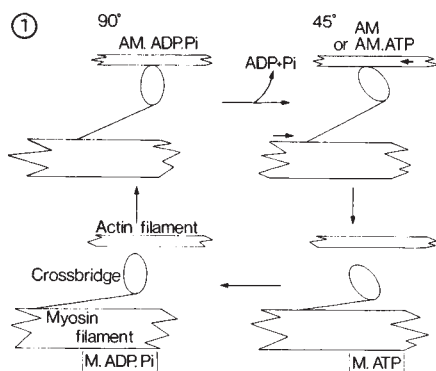
IN THE most widely accepted model of muscle contraction the actin filaments are pulled along by the cyclical action of crossbridges projecting from the myosin filaments. In this model, crossbridges are thought to attach to actin roughly at right angles to the filament axis (the so-called 90° state) and then to swing round in a power stroke that moves the actin filament by about 120 Å (refs 1,2). Binding of ATP is believed to cause detachment of the crossbridge; the ATP has to be hydrolysed before the crossbridge can rebind to actin and repeat the cycle³. This model was an attractively simple explanation of mechanochemical coupling in muscle. But it is now being suggested that crossbridge detachment is not an obligatory consequence of ATP binding⁴, that the idea of the 90° state needs major revision^{4,5} and that the power stroke is much shorter than 120 Å (ref.5). In contrast, T. Yanagida *et al.* elsewhere in this issue⁶ present an argument for a power stroke larger than 600 Å (see box). Where does this leave the crossbridge theory?

A cartoon of the original model is shown in Fig.1, where the crossbridge is drawn as a 'ball-on-a-stick'. The elongated ball is the 'head' domain of myosin (M), containing the ATP and actin binding sites. Starting at the lower left of the reaction cycle, a head with bound ADP and phosphate (P_i) binds to actin (A) to produce the complex AM.ADP.P_i. Release of the products of ATP hydrolysis is coupled with the power stroke as the head swings to 45° (AM). ATP binding detaches the head (M.ATP) from actin, the ATP is hydrolysed and the resulting complex, M.ADP.P_i, can rebind to actin in the 90° configuration. This biochemical scheme was suggested by studies in solution on the interaction of actin with isolated myosin heads³.

E. Eisenberg and T.L. Hill⁴ have recently reviewed these solution studies in the light of subsequent experiments using a wider range of conditions; it is now clear that, in solution, ATP binding to the head does not necessarily cause it to dissociate from actin before the ATP is hydrolysed. Rather, when either ATP or ADP and P_i are on the head it is weakly bound to actin, in a rapid equilibrium between attachment and detachment. A simple crossbridge cycle suggested by these newer data is shown in Fig. 2. Here there is no structural difference between the weakly bound AM.ATP and AM.ADP.P_i complexes, which have been depicted with a range of orientations to suggest the possibility that they do not have a sterically unique conformation. The power stroke is still associated with P_i release, and drives

the head into the strongly bound AM.ADP and AM states. The difference is that ATP binding now returns the head to the weak binding states rather than detaching it.

At first sight the new model seems to have a fundamental defect: the power stroke is reversed while the heads are still



attached, which must push the actin filaments the wrong way. Why does this not reverse the work done in the power stroke? Recall that a myosin filament has many crossbridges working in parallel, but asynchronously, so that swinging of an individual head will not necessarily be accompanied by movement of the actin

filament to which it is attached. The difference must be taken up as conformational strain within the attached bridge. To make the model work, Eisenberg and Hill propose that those heads in the population that are strained the 'wrong' way detach more quickly. The general idea of a strain-dependent detachment rate is not new⁷ and, indeed, is necessary for the efficient operation of any model of this general type, including that in Fig. 1. Such models can provide a quantitative description of the mechanical and energetic properties of muscle^{7,9}, but are not critically dependent on the chemical identity of the intermediates in the crossbridge cycle.

Does the biochemical cycle in Fig.2, inferred from solution studies, actually occur in contracting muscle? The answer will require biochemical studies of intact muscle fibres, in which it is difficult to apply the transient kinetic techniques used in solution. Fortunately, it has recently become possible rapidly to release ATP from a photolysable precursor inside muscle fibres¹⁰, providing at least a partial solution to the technical problem.

From the structural viewpoint the model in Fig. 2 is rather poorly defined. One general problem is that if strain in a crossbridge affects its structure then a given biochemical state may not be identifiable with a unique conformation. Nevertheless, there still might be two types of conformation corresponding to the strongly and weakly bound states. H.E. Huxley and M. Kress⁵ have now summarized the ultrastructural evidence from contracting muscle and demonstrate

A crossbridge too far ?

AN interesting new view of the crossbridge power stroke is presented by T. Yanagida, T. Arata and F. Oosawa on page 366 of this issue. They have estimated the length of the power stroke from the maximum velocity of filament sliding in crab muscle, in which the sarcomeres are long enough for the velocity of fluorescently labelled actin filaments to be measured with a light microscope. The normal transverse connection between adjacent actin filaments — the Z line — was removed by selective proteolysis, leaving single sarcomeres with a normal array of myosin filaments but with actin filaments that were now free to move independently. On addition of ATP in the presence of calcium, all the actin filaments started to slide with a velocity of about 5 $\mu\text{m s}^{-1}$. The rate of ATP hydrolysis in the same conditions was about 1 s^{-1} per myosin molecule, which corresponds to about 80 s^{-1} per actin filament.

Yanagida and his colleagues argue that as viscous drag would very quickly attenuate actin filament momentum, each filament must have at least one crossbridge pulling on it at any instant. So if one ATP molecule is hydrolysed per crossbridge cycle, the rate of cycling of active cross-

bridges could not be higher than 80 s^{-1} . To achieve a filament velocity of 5 $\mu\text{m s}^{-1}$ the length of the power stroke under these conditions would therefore have to be at least 5/80 μm or 625 Å.

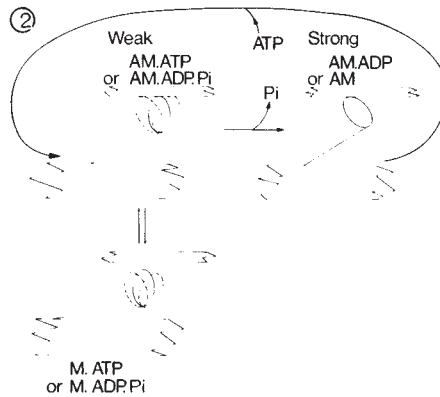
This estimate is not only strikingly larger than those obtained by other methods (see accompanying article) but also than the largest dimension of the crossbridge head (190 Å) or the actin monomer (70 Å). If the power stroke really is greater than 600 Å, some drastic rethinking of the swinging head crossbridge model would be required. The paradox might be most simply resolved, and a short power stroke retained, if the condition that one ATP molecule must be hydrolysed per cycle of crossbridge attachment and detachment did not hold during rapid filament sliding against little or no resistive force. This could occur, for example, if the number of weakly bound, rapidly attaching and detaching crossbridges increased during sliding as the degree of overlap between actin and myosin filaments became progressively greater. This might make sliding energetically favourable without committing such crossbridges to ATP hydrolysis.

Malcolm Irving

that it is consistent with the existence of two such populations.

Strongly bound crossbridges may be structurally similar to those in rigor muscle, where no ATP is present. If such heads were attached to actin in a stereospecific way and all had the same shape they would take up the periodicity of the actin filament helix. X-ray diffraction studies suggest that no more than 20 per cent of the heads are attached in this way during contraction (compared with 100 per cent in rigor). This is marginally consistent with electron spin resonance studies¹¹, which indicate that in about 20 per cent of the crossbridge heads the probed domain has the same orientation as in rigor; the rest of the heads are disordered and rapidly mobile. However, another part of the X-ray diffraction pattern shows that at least 60 per cent of the heads are near the actin filaments during contraction. Huxley and Kress suggest that in addition to the 20 per cent of strongly bound heads, a further 40 per cent are weakly bound, with a wide range of permitted orientations and in a rapid equilibrium of attachment and detachment. Independent evidence for this type of crossbridge attachment has come from studies on demembrated muscle fibres in calcium-free, low ionic strength media — conditions in which isolated heads in solution are weakly bound to actin. X-ray diffraction studies on such fibres showed that the crossbridges were near actin but had not taken up its helical periodicity¹² and, although no force is generated, mechanical studies detected a population of rapidly attaching and detaching crossbridges¹³.

Do the weakly bound crossbridges in contracting muscle also generate no force? If only the strongly bound 20 per cent of heads do so, one can calculate that each would have to exert about 6 piconewtons in a slowly shortening muscle. But if this is applied over a 120 Å power stroke the work produced per cycle (7.2×10^{20} joules) would be about three times greater than expected from ATP hydrolysis on the basis of the observed macroscopic efficiency. Huxley and Kress conclude from a similar calculation that if only 20 per cent of heads produce force the power stroke must be smaller than originally proposed—they suggest it is about 40 Å. The previously favoured value of 120 Å came from experiments in which extremely rapid shortening was imposed on active muscle²; fast force recovery was seen only if the filament sliding was less than 120 Å. Huxley and Kress suggest an alternative interpretation of these experiments in which 120 Å becomes the range of attachment of the weakly bound heads. Although these would not normally be force producing, filament sliding of up to 120 Å would bring them into the force-generating zone where they could make the transition to the strongly bound state. A 40 Å power stroke of a strongly bound head could plausibly occur in an all-or-



none fashion, storing energy in the conformational strain of another part of the crossbridge. This would explain why the ordered heads in contracting muscle have exactly the rigor conformation (as detected by spin-resonance probes): strongly bound heads are only seen at the end of their power stroke.

Many questions remain. Much of the mechanical and structural data could be explained without postulating the existence of weakly bound, non-force-generating crossbridges in contracting muscle. In an alternative view most of the heads would be strongly bound, force generating and take up a wide range of conformations with rapid transitions between them². A key proposal of the model of Huxley and Kress, that only 20 per cent of heads are strongly bound and force generating, derives from studies with extrinsic probe molecules. It will be important to check by independent methods that crossbridge properties have not been modified by the probes and that the mobil-

ity of the probe does indeed reflect that of the head. This proposal also seems difficult to reconcile with existing data from experiments designed to compare the mechanical properties of contracting with rigor muscle¹⁴ which suggest that the fraction of force-generating heads during contraction may be as high as 75 per cent. Nevertheless, the idea that there are two types of crossbridge in contracting muscle—weakly and strongly bound—provides a neat and stimulating synthesis of a wide range of muscle phenomena. In the longer term its success will depend on the extent to which these two hypothetical states stand up to attempts at making a more rigorous characterization of their structural and dynamic properties. □

1. Reedy, M.K., Holmes, K.C. & Tregear, R.T. *Nature* **207**, 1276 (1965).
2. Huxley, A.F. & Simmons, R.M. *Nature* **233**, 533 (1971).
3. Lynn, R.W. & Taylor, E.W. *Biochemistry* **10**, 4617 (1971).
4. Eisenberg, E. & Hill, T.L. *Science* **227**, 999 (1985).
5. Huxley, H.E. & Kress, M.J. *Mus. Res. Cell Motil.* **6**, 153 (1985).
6. Yanagida, T., Arata, T. & Oosawa, F. *Nature* **316**, 366 (1985).
7. Huxley, A.F. *Prog. Biophys. biophys. Chem.* **7**, 225 (1957).
8. Hill, T.L. *Prog. Biophys. molec. Biol.* **28**, 267 (1974).
9. Eisenberg, E., Hill, T.L. & Chen, Y. *Biophys. J.* **29**, 105 (1975).
10. Goldman, Y.E., Hibberd, M.G., McCray, J.A. & Trentham, D.R. *Nature* **300**, 701 (1982).
11. Cooke, R., Crowder, M.S. & Thomas, D.D. *Nature* **300**, 776 (1982).
12. Matsuda, T. & Podolsky, R.J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2364 (1984).
13. Brenner, B., Schoenberg, M., Chalovich, J.M., Greene, L. & Eisenberg, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7288 (1982).
14. Goldman, Y.E. & Simmons, R.M. *J. Physiol., Lond.* **269**, 55P (1977).

Malcolm Irving is in the Department of Biophysics, King's College London, 26–29 Drury Lane, London WC2B 5RL, UK.

Structural geology

To stretch a continent

from I. R. Vann

THAT large areas of the Earth's crust have been subjected to compression has been known since before the discovery of plate tectonics. But, outside the world of hydrocarbon exploration, extensional tectonics has been a topic of intensive study for less than a decade. A recent conference* illustrated the remarkable progress that has arisen from two very different scientific approaches, the first a deduction straight from the world of physics, the other an induction rooted firmly in the geological tradition of careful field observation.

Seven years ago, D. McKenzie made what in retrospect seems a trivially simple prediction (*Earth. planet. Sci. Lett.* **40**, 25; 1978) — if the lithosphere is stretched then it must thin. Since the base of the lithosphere is an isotherm, a thermal anomaly is created that slowly decays causing subsidence of the crust as an inverse exponential function of time. The

power of this model is that it connects the stratigraphy and thermal history of the sediments filling a basin with a regional structure. At around the same time, field studies in the 'Basin and Range', the core of the Rocky Mountains of the western United States, began to show that the structure of the area is dominated by flat-lying extensional faults (detachments) with horizontal offsets across them of tens of kilometres.

This observation dispelled an old geological prejudice that sub-horizontal faults are all compressional structures (thrusts). Despite this a number of questions remain, concerning the original geometry of the flat detachments, the relevance of the Basin and Range model to other extensional basins and the relationship between the extension recorded near to the surface in rift basins and deeper structures.

At the conference the problems of unravelling the original geometries of flat detachments was exemplified by the con-

*'Continental Extensional Tectonics', University of Durham, UK, 18–20 April 1985.

tributions of G. H. Davis (University of Arizona), and E. L. Miller (Stanford University) and her co-workers. Davis suggested that the faults and the 'metamorphic core complexes' that underlie them were originally moderately dipping ductile zones of simple shear formed at 10 km or deeper within the crust. As extension progressed, the material overlying those 'faults' was moved laterally by tens of kilometres. The overall effect of this is to bring the faults nearer the surface and thus reduce the confining pressure and temperature. This causes a change in the deformation from a ductile to an increasingly brittle process. According to Davis, younger cross-cutting extension faults later rotated the dipping zones to horizontal.

E. L. Miller's model is markedly different. She suggested that the Snake Range detachments in Nevada are neither faults nor shear zones but the transition, formed at a depth of 12 km, between an upper brittle plate and a lower ductile one, which latter deformed by pure shear expressed as horizontal extension and vertical thinning. Many other speakers accepted in modified form one or other of these apparently incompatible models but it was disconcerting for the outsider to see that this marvellously exposed area remains so enigmatic.

Even if the Basin and Range ultimately yields a single unified model it will remain a unique area of extreme extension within a region of relatively recent crustal thickening in the core of a thrust mountain belt. A possible ancient analogue — the Devonian (380 Myr old) basins of western Norway — was described by M. Seranem, M. Seguret and Ph. Laurent (University of Montpellier II). Here 50 km or more of extension occurs above a shear zone, apparently a reactivated thrust, which dips westwards at 5° to 25°, thereby creating a basin containing sediments up to 25 km thick. Geometrically similar basins, observed on deep seismic profiles (Brewer, J. A. & Smythe, D. K. *J. geol. Soc. Lond.* **141**, 105; 1984), occur above similarly reactivated thrusts on the western margin of the Caledonian mountain belt in north-west Scotland, which appear to flatten as they enter the lower crust at a depth of around 18 km. These fragmentary observations suggest that an extensional province similar to the Basin and Range existed within the Caledonian mountains some 380 Myr ago.

Is it possible to carry the Basin and Range analogy further? Certainly there are great differences between the extension within mountain belts and that which creates great sedimentary basins (rifts) in old stable crust (such as the East African Rift). Perhaps the most obvious of these differences is that almost all these rifts are covered by water or thick blankets of sediment. Hence their structure is not accessible to the geologist's hammer and our understanding depends largely on seismic

reflection data that are even more ambiguous than surface observation. A. D. Gibbs (Midland Valley Exploration, Glasgow) presented models of rift fault systems developed in the North Sea that exemplify the belief of a growing number of workers that extension is achieved by movement on a linked series of faults that have flatter and steeper portions producing a 'staircase' geometry in the upper crust. These faults systems are strongly asymmetric giving rise to basins bounded on one side only by major faults. Along the length of any basin this asymmetry flips from side to side across complex transfer zones.

Unfortunately the seismic data, which has been acquired almost exclusively by oil companies, images only the upper 10 km of the crust. Basin and Range models are applicable, if at all, to rift systems at depths of more than 15 km where it is thought that the crust changes from a generally brittle to a ductile material. Evidence about deformation at these depths below rifts comes from a very few deep reflection experiments conducted mainly on the UK continental shelf. A. Beach (Britoil, Glasgow) presented one such profile across the northern North Sea. His interpretation suggests two major periods of extension in this rift with active faulting in Triassic and early Jurassic times (220–200 Myr) above a single low-angle shear zone akin to the Basin and Range model

of G. Davis being followed by extension at the end of the Jurassic along steep faults that cut entirely through the crust. The profile demonstrates that the crust below the rift is only half as thick as it is on the flanks and hence supplies one of the few direct tests of the McKenzie stretching model.

Support for Miller's interpretation of the Basin and Range came from a discussion by J. A. Jackson (Cambridge University) of deep deformation processes within active regions of extension. He noted that all major earthquakes in active regions occur on faults dipping at 30–60° at depths of 8–12 km and are grossly planar from the surface to these depths. Below this earthquakes are rare, but when they do occur (usually as aftershocks to large shallower events) they seem to do so on horizontal fault planes. Jackson concludes that extension in the lower crust normally takes place by ductile diffuse creep with only minor brittle effects.

Despite unanswered fundamental questions, and amidst much conflicting evidence and irreconcilable theory, many participants left the conference feeling that they had attended an important event. Among the conflict it seems there is a convergence that in a very few years will lead to a unifying 'general theory' of extensional tectonics. □

I. R. Vann, is a geologist with BP Petroleum Development, Britannic House, Moor Lane, London, EC2Y 9BU, UK.

Infrared astronomy

Views across the Galaxy

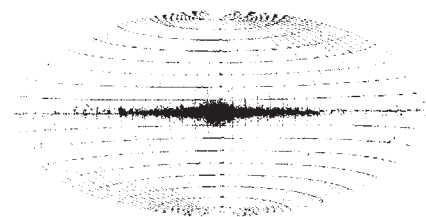
from Charles Beichman

No longer the coveted property of just a few infrared astronomers, the wealth of data returned from the Infrared Astronomical Satellite (IRAS) is now being studied by a large number of scientists working in all regions of the spectrum. Three days of a recent conference* were devoted to the presentation and discussion of IRAS discoveries within our Galaxy. More than a dozen invited talks and fifty contributed papers described what IRAS has revealed about the death and birth of stars, the large-scale structure of the Galaxy and properties of interstellar dust.

An examination of the overall statistics of the IRAS point source catalogue (T. Chester, California Institute of Technology) demonstrates the importance of IRAS to galactic astronomy. Half of the 245,839 sources detected by IRAS lie within seven degrees of the galactic plane; two thirds of the sources lie within the central two quadrants of the Galaxy. Roughly 158,000 stars within the Galaxy comprise 65 per cent of the catalogue. Some 65,000 objects are associated with different phases of the interstellar

medium, including H II regions, molecular clouds with embedded pre-main sequence stars and dense clumps of cold dust. The remaining 9 per cent of the catalogue consists of about 22,000 external galaxies. Spectroscopic data obtained with the IRAS Low Resolution Spectrometer are available for the 5,000 brightest objects (F. Olmon, Groningen).

At the end of its life, a normal star ejects a large part of its mass back into interstellar space. The resulting shells of circumstellar dust can convert most of a star's luminosity into infrared energy and make such stars detectable by IRAS across nearly the entire Galaxy. H. Habing (Leiden) used stars with shells detected at 12 and 25 microns to trace the disk and central bulge of the Galaxy with unprecedented clarity (see figure), free from the effects of inter-



*The IRAS Symposium "Light on Dark Matter" was held at Noordwijk aan Zee, the Netherlands, 10–14 June 1985.

stellar extinction that preclude such studies at visual wavelengths. The colours and luminosity of these stars resemble those of previously known OH/IR stars, which, in addition to being bright in the infrared, emit intense radiofrequency radiation in the spectral lines of the OH radical. Ground-based observations (J. Frogel, Kitt Peak; M. Feast, South African Observatory) confirm that many IRAS stars found in the bulge are late-type M giant stars, with luminosities as large as 10^3 to $10^4 L_{\odot}$, that are losing mass at a rate of 10^{-6} to 10^{-4} solar masses per year. The age of these stars and the mass of their main sequence progenitors are quite uncertain, but an intriguing possibility is that they were originally one-solar-mass stars and are no more than a few billion years old, far less than the age of the Galaxy. In that case, there may have been a burst of star formation in the bulge of the Galaxy that resembled bursts seen towards other infrared-bright galaxies detected by IRAS.

B. Zuckerman (University of California, Los Angeles) gave two reasons why an understanding of the mass loss in these late-type stars is important. First, as we observe many white dwarf stars and relatively few supernovae in the Galaxy, mass loss must provide the mechanism whereby stars more massive than 1.4 solar masses lose enough material to avoid a catastrophic demise when their nuclear fuel has been consumed. Second, it is through mass loss that the carbon, nitrogen and oxygen produced by nuclear burning are injected into the interstellar medium for subsequent processing into new stars, planets and, eventually, astronomers. By using a combination of IRAS measurements and millimetre-wavelength observations of gas outflows for a large sample of IRAS stars, it will be possible to learn much about the dynamics and chemistry of the mass-loss process and its relation to later stages of stellar evolution.

P. Bedijn (Heidelberg) explained the cause of mass loss in terms of a positive feedback mechanism coupled to global pulsations of the star. His theory derived support from the fact that more than half of the IRAS sources varied in brightness by as much as a factor of three over the course of the IRAS mission.

After the mass-loss phase ends, it is thought that late-type giants become planetary nebulae which consist of a 50,000 K star surrounded by a thin shell of ionized gas. A. Leene and S. Pottasch (Leiden) used 8 to 22-micron spectra obtained with the Low Resolution Spectrometer to show that for some of the more than 700 planetaries detected by IRAS, the radiation does not come from hot dust, but rather from spectral lines from ionized gas. Emission from spectral lines of S IV (triply ionized sulphur), Ne II, O IV and S III account for all of the radiation detected by IRAS in the 12- and 25-micron bands of the survey. Leene and

Pottasch predicted that lines of N II and O III could account for some or all of the 60- and 100-micron radiation detected in the survey for these objects. IRAS maps of planetary nebulae can be used to examine the temperature, abundance, density and excitation of these ionized species.

IRAS is also well suited to studying the other extreme of the life history of stars — their birth out of the interstellar medium. The physics of star formation represents a fundamental problem in modern astronomy which is relevant both to the structure of our own Galaxy as well as to the energy production mechanisms for the most luminous starburst galaxies detected by IRAS. As B. Elmegreen (IBM, T.J. Watson Center) pointed out, IRAS data at the highest sensitivity and spatial resolution can be used to study the detailed physics of star formation, and IRAS images can be used to study large-scale processes in the interstellar medium that produce stars.

One mechanism for the formation of stars that has been discussed for many years is that H II regions around newly formed massive O and B stars trigger the fragmentation and subsequent collapse into new stars of adjacent material in a giant molecular cloud. By examining the 60- and 100-micron images produced by IRAS, P. Schwartz (US Naval Research Laboratory) found circumstantial evidence for loops and rings of star formation surrounding earlier sites of star formation. One such loop is 250 pc in diameter and contains many H II regions and embedded young stars. If further examination of the IRAS images shows that star formation is often associated with structures of this type, then the case for triggered or sequential star formation will be greatly bolstered. As Schwartz emphasized, large-scale morphological studies of this kind are possible only with the advent of the IRAS images that present to the eye in a single picture both the distribution of the

interstellar matter and the presence of embedded young stars.

A second way by which star formation may proceed is to compress small globules of interstellar material enough to overcome the supporting motions within the cloud and so initiate a gravitational implosion that leads to the formation of one or more stars. It is possible that many solar-type stars form in this manner, rather than in the triggered fashion of the more massive stars. The existence of this more sedate and spontaneous type of star formation has been demonstrated by P. Myers (Center for Astrophysics), C. Beichman (California Institute of Technology), S. Harris and J. Emerson (Queen Mary College, London) and R. Jennings (University College, London) who examined more than 100 small clumps of interstellar gas in the Taurus and Ophiuchus molecular clouds, some 150 pc away. These clumps contain between 5 and 50 solar masses of material and represent the simplest possible sites for star formation. The IRAS survey shows that about half of these clouds are associated with infrared point sources. Although a few objects can be identified with visible pre-main sequence stars (T Tauri stars), the remainder have no visual counterparts. The invisible sources have very low colour temperatures, about 100 K, and luminosities of 0.5–10 $L_{\odot}$. Although it is not possible to determine the masses of such objects accurately, it is likely that stars of about 0.5–1 solar mass are being formed in these clumps.

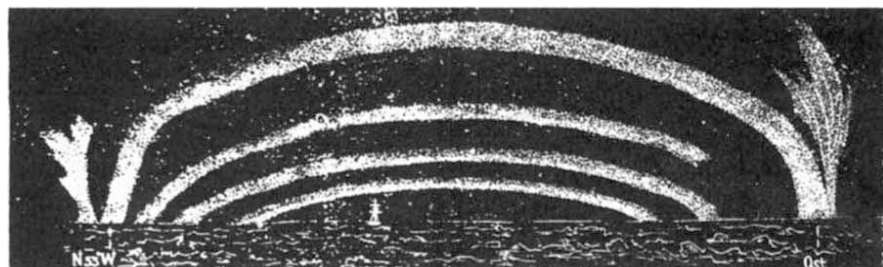
The most important, but as yet unanswered, question about these embedded infrared sources concerns their evolutionary status. Some of the invisible infrared objects are probably T Tauri-like stars, no older than 10^5 years, that obtain their energy from the slow contraction of a stellar core in quasi-static equilibrium, but which are still obscured from view by placental gas and dust. On statistical

100 years ago

MR TROMHOLT has rendered a great service to science by the travels and observations recorded in these volumes¹; indeed, it would not perhaps be going too far to say that we have here, brought before us in the most interesting manner, one of the best organised attempts to study the aurora that has been made for many years, the credit for which must be given to the organisers of the International Polar Research Expedition of 1882–83. The results, however, recorded in these volumes are by no means limited to the height of the aurora. The con-

THE AURORA

stant study afforded to Mr. Tromholt of one of the most awe-inspiring phenomena which it is given to man to witness have permitted generalisations to be reached and hypotheses to be broached of the greatest scientific interest. Mr Tromholt holds that *the auroral zone moves northwards and southwards daily, yearly, and eleven yearly*. To speak in vacuum-tube language, instead of one rigid stria, we may have many striæ, and these moving towards or away from the auroral pole as ordinary striæ move towards or away from the negative pole.



Edited by Carl Siewers. (London: Sampson Low & Co., 1885.)

From *Nature* 37 274, 25 July 1885.

grounds, however, it is probable that some of these IRAS sources are still deriving most of their luminosity from the release of the gravitational energy of infalling cloud material. These objects can be considered to be true protostars. Confirmation of this status and further study must rely on observations at other wavelengths, in particular at high spatial and spectral resolution in the near-infrared and at millimetre wavelengths.

Infrared astronomy can be considered to be the study of dust in one form or another, and a large part of the conference reflected the importance of solid material, heated enough to be detectable in the infrared. The so-called infrared cirrus, diffuse emission from interstellar dust, is an important and puzzling discovery of IRAS. T. N. Gautier (California Institute of Technology) and others showed that at 100 microns the cirrus can be correlated with faint dust features visible on high-contrast optical photographs and with structures seen in 21-cm maps of atomic hydrogen (H I). Some cirrus can be associated with weak emission seen in high galactic molecular clouds.

Some of the most exciting results concern the diffuse emission seen at 12 microns by IRAS. The properties of 12-micron cirrus tend to confirm the work of K. Sellgren (Hawaii) that there exists dust whose apparent colour temperature is much hotter than its equilibrium temperature. The explanation of this effect is that motes of dust a few tens of angstroms in size can be heated, briefly, by the absorption of a single ultraviolet photon. The heated grains cool off by emitting radiation at anomalously hot temperatures. A. Leger, J. L. Puget and F. Boulanger (Paris) have identified polycyclic aromatic hydrocarbons as a likely type of grain material. Composed of carbon and hyd-

rogen ions, they are not only stable at the small sizes required, but also have spectral features at 3 and 11 microns that match previously unidentified bands found towards several infrared sources.

The realization that very small grains play an important role in the interstellar medium was brought out repeatedly during the conference. Elmegreen pointed out that if very small grains exist at the centres of dense molecular clouds and not just at their periphery, then they will be charged and will control the coupling of the gas to the magnetic field. The properties and abundance of these grains may control the collapse of molecular clouds.

The 60- and 100-micron IRAS maps have been used to examine the large-scale structure of the interstellar matter in the Galaxy. B. Burton, H. Walker and E. Deul (Leiden) found that the scale height of the 100-micron emission is about 120 pc, the same as that of the atomic hydrogen (H I), and twice as great as the scale height of the molecular gas as inferred from CO observations. Thus, IRAS has traced the distribution of warm (22 K) dust in relatively diffuse regions and not the distribution of the colder, denser material. A longer wavelength map of the Galaxy, say at 200–300 microns, would trace the 10–15 K material associated with the molecular component of the Galaxy.

The IRAS catalogue and all sky images have been generally available only since last December. The number and variety of results reported at the conference suggest that the data returned by IRAS will be a valuable resource for astronomers for years to come. □

Charles Beichman is at the Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, California 91125, USA.

the differences in relative numbers will manifest themselves as changes with time in the integrated properties of the burst of star formation ('starburst').

There are several ways in which the efficiency of star formation may be related to large scale morphological properties of galaxies—another fundamental problem. For example, T. Harwarden (Royal Observatory, Edinburgh) suggested that the presence of galactic bars may provide a kinematic catalyst which channels material inward to sites of galactic nuclear starbursts. An obstacle to progress in this area is the sheer difficulty of quantifying objectively the pictorial evidence of bar structure. Another process that may enhance the efficiency of, and/or provide a trigger mechanism for, a burst of star formation is that of galaxy interactions. R. Joseph and co-workers (Imperial College, London) have surveyed interacting pairs of galaxies and find evidence for starbursts to be very common, although characteristically only in one member of a pair.

The key observational inputs to star formation models were discussed by C. Telesco (NASA Marshall Space Flight Center). An important parameter is the total infrared luminosity of a star-forming region which, roughly speaking, indicates the number of young stars present. For this purpose the IRAS observations are ideal, covering as they do a range of 12 to 100 microns. Another global property is the mass/light ratio. In the conventional units normalized to the solar values it is known that an old stellar population has a ratio of greater than three, whereas it is predicted that the ratio for an unevolved stellar population (formed with the mass distribution observed in the solar neighbourhood) is less than 0.01. This index can therefore be used as a measure of the age of the stellar population and hence of the presence and strength of a starburst. Telesco described the application of a starburst model to the famous galaxy Messier 82. The implied age of the starburst in M82 turns out to be 20 million years. Using the far infrared luminosity, the deduced rate of star formation is 34 solar masses per year which, integrated over the age of the starburst, is a large fraction of the total mass. To avoid the situation in which such a large fraction of all mass has been converted, it is necessary to impose a cut-off in the mass distribution of the new stars. Thus, in luminous starbursts there would be relatively few new low-mass stars, high-mass conversion rates and short-burst durations. These findings will have important implications for models which follow the evolution of starbursts.

The blue compact galaxies, also known as extragalactic H II regions, are relatively unevolved stellar systems. Their near-infrared colours indicate that they are young and/or metal-deficient, meaning that relatively little mass has been converted into heavier elements via stellar evolution. In support of this, IRAS obser-

Infrared astronomy

Views outside the Galaxy

from Martin Ward

ALTHOUGH the infrared satellite mission IRAS ended its life more than a year and half ago, its influence on the subject of infrared astronomy promises to continue for many years to come, by virtue of the enormous database of observations that it produced. That this is so for extragalactic astronomy was demonstrated at a recent workshop*, the main thrust of which was aimed at the understanding of aspects of star formation within galaxies.

Young stars are hot and emit most of their energy in the optical and ultraviolet regions of the spectrum. But as dust and gas exist together in the star-forming regions, some of these optical and ultraviolet photons heat up the dust grains,

which then radiate approximately like black bodies. The exact temperature of the heated dust will depend on various factors but, leaving aside the more exotic active galactic nuclei for the moment, the range extends from 30 to 100 K. These temperatures produce peak emission in the infrared, hence the connection between star formation and far-infrared emission.

As was emphasized by H. Yorke (University of Göttingen), there are still fundamental questions that need to be answered concerning star formation and the evolution of H II regions. For example, relatively how many new stars of various masses are formed in any one episode of star formation? Since high mass stars evolve faster than those of low mass,

* The meeting on Extragalactic Infrared Astronomy was held at the Rutherford Appleton Laboratory, 20–27 May 1985.

vations show that the gas/dust ratio in such galaxies is a thousand times higher than that in the solar neighbourhood (P. Gondhalekar, Rutherford Appleton Laboratory); the most probable implication is that there is a pronounced paucity of dust. If dust is formed in stellar atmospheres, the low dust content implies either a slow rate or late start of star formation. It is still controversial as to whether we are witnessing the initial starburst or whether these galaxies have experienced a series of previous bursts.

Naturally the more exotic examples of active galaxies also received some attention. I reported new infrared results for an X-ray selected sample of Seyfert galaxies. Although there are exceptions, in most cases the dominant mechanism responsible for the strong infrared emission is not a massive burst of star formation such as observed in M82. It seems rather that the near infrared emission at a few microns is related, but not necessarily via the same physical process, to emission from the very compact regions. Good correlations are found with the strengths of the very broad emission lines and the X-ray emission, whilst at 20 microns the best correlation is with the narrow emission lines which originate in a volume of typically a thousand light years. This 20-micron luminosity could come from warm dust (200 K) mixed in with the line-emitting gas and heated by the very compact source at the centre of the active nucleus.

Because of the large apertures of the IRAS observations, their relation to the compact active nuclei in Seyfert galaxies is not clear. It seems probable that we are seeing a 'mixture' of several spatially-distinct components. Michael Rowan-Robinson (Queen Mary College, London) showed how, for Seyfert galaxies, the IRAS data can be separated into three

components: a cooler disc component (seen in more normal galaxies), a warmer starburst component and a Seyfert component peaking at 25 microns. The latter could be radiation from a mini-quasar at the centre of the Seyfert absorbed by dust and re-radiated in the infrared.

The cosmological implications of the IRAS observations were also considered by Rowan-Robinson. Concentrating on the galactic polar caps, where confusing effects of the infrared 'cirrus' (see *Nature News and Views* 308, 224; 1984) are smallest, he showed that there is a significant anisotropy in the number density of IRAS 60-micron sources between the north and south polar caps.

This seems not to be a local phenomenon since an identification programme carried out for the northern cap reveals that, after elimination of obvious stars, virtually every source is a galaxy and that the typical distance of IRAS galaxies is about 200 megaparsecs. Assuming that the infrared radiation traces the matter, it is possible to map the gravitational field and determine its dipole component. The direction of this component is in good agreement with the observed anisotropy in the intensity of the cosmic microwave background radiation. Thus this anisotropy could be explained as the result of the net gravitational pull on our local group of galaxies by the distribution of galaxies within about 200 megaparsecs.

Rowan-Robinson also showed that the 100-micron background radiation detected by IRAS can be separated into components due to interplanetary dust, interstellar dust and a roughly isotropic component which may be of cosmological origin. □

Martin Ward is at the Institute of Astronomy, Madingley Road, Cambridge CB3 0HE.

Solid-state chemistry

Topotaxy in metal-oxide reduction

from Robert W. Cahn

A THIN metal film deposited by evaporation or sputtering on, say, a cleavage surface of sodium chloride is apt to be oriented in a reproducible way with respect to the substrate — this is the commonly found physical process of epitaxy, where the substrate is inert and unaffected. The chemical analogue is a topotactic reaction¹: here a parent phase undergoes a reaction and the product phase has one or more defined orientations relative to the parent. Such reactions have been analysed by metallurgists (although they never use the term topotaxy) ever since the discovery of X-ray diffraction: the key investigations on so-called Widmanstätten reactions in metallic solid solutions were carried out in the early 1930s. Metallurgists usually analyse

such reactions in terms of the fit of parent and product phases at their mutual interface.

The reverse reaction, reduction of an oxide to metal, has occasionally been examined from a kinetic point of view, especially in connection with the industrially important reduction of wuestite (FeO) to iron², but the first study of topochemical aspects has only just been performed. The results, reported by A. Revcolevschi and G. Dhalenne on page 335 of this issue³, open up a new field of research.

In chemistry proper, topotaxy is usually demonstrated in dehydration reactions — for example, the conversion of Mg(OH)₂ into Mg(O) — or the conversion of one oxide of iron into another; other instances include breakdown of complex silicates

into simpler ones, with release of water⁴. In all these reactions, the observed orientations imply that the network of oxygen ions remains more or less intact throughout the course of the chemical reaction.

The most studied category of topotactic reaction has been the oxidation of metals. The nature of both the topotaxy of a thin oxide layer and the oxidation kinetics depends not only on the oxidation conditions but also on the orientation of the metal surface (see refs 5 and 6 for review).

Revcolevschi and his collaborators at the University of Paris-Sud have for some time studied the orientation relationships between the phases of lamellar eutectic structures consisting of two metallic oxides (complementing the substantial body of information on such relationships as regards metallic eutectics⁵). For their new experiments, Revcolevschi and Dhalenne³ have directionally solidified a eutectic of NiO and calcia-stabilized cubic ZrO₂, producing what is essentially an interpenetrating pair of epitaxially related single crystals which they then treat with a CO/CO₂ mixture that reduces the NiO to nickel without affecting the ZrO₂. The nickel proved to be in parallel orientation to the NiO which preceded it; the ZrO₂ phase acted as crystallographic reference. Just as earlier topotactic experiments indicated that the oxygen skeleton was maintained and unchanged, so the new experiment³ indicates that the nickel-ion skeleton remains fixed.

This pioneering experiment opens the way to various topotactic studies. One might try to mimic Gwathmey's celebrated experiment on the oxidation of a spherical copper monocrystal (showing the sharp anisotropy of oxidation rates and topotaxy) by reducing an oxide sphere, or to study the orientation dependence of reduction kinetics. It is known that the reduction rate of wuestite is sensitive to the population of point defects — it would be interesting to see whether orientation relationships are affected by defects.

Revcolevschi and Dhalenne's observations have a potential practical interest: by reducing NiO without altering ZrO₂, it is possible to make a lamellar product where adjacent lamellae are electronically and ionically conducting. Such a structure can perhaps be applied in photoelectrolysis and novel solid-state batteries. □

1. Thomas, J. M. *Phil. Trans. R. Soc. A* 227, 251 (1974).
2. Schmalzried, H. *Solid State Reactions* 195 (Academic, New York, 1974).
3. Revcolevschi, A. & Dhalenne, G. *Nature* 316, 335 (1985).
4. Brindley, G. W. *Prog. Ceram. Sci.* 3, 1 (1963).
5. Gwathmey, A. T. & Lawless, K. R. in *The Surface Chemistry of Metals and Semiconductors* (ed. Gatos, H. C.) 483 (Wiley, New York, 1960).
6. Bardollé, J. in *Interfaces et surfaces en Metallurgie* (eds Martin, G. et al.) 467 (Trans Tech, Aedermannsdorf, 1975).
7. Chadwick, G. A. *Prog. Mat. Sci.* 12, 99 (1963).

Robert W. Cahn, formerly Professor of Materials Science at the University of Sussex, is currently Visiting Research Fellow at the General Electric R&D Center, Schenectady, New York 12301, USA.

Plant ecology

Fertile arguments in the desert

from Jonathan Silvertown

ECOLOGISTS have been arguing recently about whether interspecific competition is responsible for observed patterns of animal distribution and abundance^{1,2}. Whereas interspecific competition once seemed to be the only process of interest to the theoretically minded, now dispersal, predation, mutualism and biogeographic history are also being considered (see ref. 1). Strangely, the recent debate has occurred almost exclusively among animal ecologists but similar arguments among plant ecologists now seem to be approaching the same conclusion.

Plant ecologists studying species distri-



Larrea bushes in the Chihuaha desert.

bution in deserts have investigated the cause of the regular spatial patterns of distribution often found in the southwestern deserts of North America and elsewhere. The commonest shrub in the Chihuaha, Sonora and Mojave deserts is the creosote bush *Larrea tridentata*, which is a conspicuous part of the sparse vegetation. Bushes, or clumps of bushes, commonly have large areas of bare ground between them and often appear regularly spaced. F. W. Went suggested that this regular spacing is the result of allelopathic (toxic) interactions between bushes³. Since then, in one of the first quantitative analyses of the dispersion pattern of *Larrea*, S. R. J. Woodell *et al.*⁴ showed that *Larrea* bushes are indeed regularly spaced at sites of low rainfall but are aggregated at sites of high rainfall. In an alternative hypothesis they suggested that the regularity is the result of competition for water between roots of adjacent bushes, with reduced competition at sites of high rainfall allowing aggregation.

M. G. Barbour^{5,6} found that regular patterns of *Larrea* were in fact quite rare. D. J. Anderson⁷ further questioned Woodell *et al.*'s claim that plants in lower rainfall areas compete for water, because bushes at these sites are at such low density that each plant actually has more available soil water than in high rainfall areas where bushes are more crowded. T. J. King and Woodell⁸ replied that the quantity of water available to bushes cannot be calculated simply from density and rainfall measurements because plant size,

surface-water runoff and infiltration of water into the ground also has to be considered. They also suggested that periods of severe drought in the past might at some sites be responsible for present-day distributions. Thus, in defending the competition hypothesis, they acknowledge the importance of other factors in structuring *Larrea* communities, as animal ecologists have had to do in defence of similar arguments about interspecific patterns⁹.

The parallel with the debate among animal ecologists can be taken further. Reassessment of competition in animal ecology has been stimulated by the use of 'null models'² that predict how community structure would appear in the absence of competitive interactions between species. T. A. Ebert and G. S. McMaster¹⁰ have used this approach to challenge the regularity of spacing patterns in *Larrea*, which they believe may be no more than an artefact of sampling methods, by which plants that grow very near one another are recorded as a single individual. They show that this can cause a random pattern to appear erroneously as a regular one. King and Woodell¹¹ dispute this claim, but do not explain how the competition hypothesis accounts for regular spaced clumps.

A general conclusion from the use of null models in animal ecology has been that patterns can only provide correlative evidence of the processes that structure communities. Experiments are needed to test hypotheses generated by analysis of pattern distribution. Several relevant experiments have now been performed with desert shrubs. In the field, removing neighbours of individuals of both *Larrea*

and another desert shrub, *Ambrosia dumosa*, affects the physiological water status of other plants^{12,13}. Even though this is evidence of competition for water, some of the experimental populations were aggregated. Clearly, competition does not lead inevitably to regular spacing.

We can say more than this. The desert perennial *Eriogonum inflatum* with competitors nearby is more likely to die than those plants with more distant neighbours, yet this does not change the degree of aggregation in the population. The most elaborate recent experiment¹⁴ shows how uncertain the role of root competition in producing regular spacing has become. Despite a fall in shrub density, W. H. Schlesinger and C. S. Jones found no change in the dispersion pattern of *Larrea* and *Ambrosia* at a Mojave site that has been deprived of surface water runoff by a drainage system for 45 years.

Ironically, the unknown causes of regular spacing in desert plant communities may bear some relationship to the way in which animal and plant ecologists have ignored each other. Few have considered, for example, that these plant patterns could result from the activities of seed-gathering rodents¹⁶ which forage and remove seeds between bushes and which are ubiquitous in these deserts. □

1. Strong, D. R. *et al.* *Ecological Communities. Conceptual Issues and the Evidence* (Princeton University Press, 1984).
2. Harvey, P. H. *et al.* *Rev. Ecol. Syst.* **14**, 189 (1983).
3. Went, F. W. *Scient. Am.* **192**, 68 (1955).
4. Woodell, S. R. J. *et al.* *J. Ecol.* **57**, 37 (1969).
5. Barbour, M. G. *Ecology* **50**, 678 (1969).
6. Barbour, M. G. *Am. Midl. Nat.* **89**, 41 (1973).
7. Anderson, D. J. *J. Ecol.* **59**, 555 (1971).
8. King, T. J. & Woodell, S. R. J. *J. Ecol.* **61**, 761 (1973).
9. Gilpin, M. E. & Diamond, J. M. *Oecologia* **52**, 75 (1982).
10. Ebert, T. A. & McMaster, G. S. *J. Ecol.* **69**, 559 (1981).
11. King, T. J. & Woodell, S. R. J. *J. Ecol.* **72**, 295 (1984).
12. Fonteyn, P. J. & Mahall, B. E. *J. Ecol.* **69**, 883 (1981).
13. Ehleringer, J. R. *Oecologia* **63**, 153 (1984).
14. Wright, S. J. *Oecologia* **54**, 266 (1982).
15. Schlesinger, W. & Jones, C. *Bot. Gaz.* **145**, 116 (1984).
16. Brown, J. H. *et al.* *Rev. Ecol. Syst.* **10**, 201 (1979).

Jonathan Silvertown is a Lecturer in the Biology Department, The Open University, Milton Keynes MK7 6AA, UK.

Excitation-contraction coupling

The messenger across the gap

from Andrew P. Somlyo

EXCITATION of the surface membrane of muscle results in release of calcium that is stored in an intracellular membrane system, the sarcoplasmic reticulum (SR); the Ca^{2+} released into the cytoplasm, acting through regulatory sites on Ca-binding proteins, initiates contraction. The signal from the membrane is communicated to the SR at specialized regions (triads, diads or surface couplings) where the cytoplasmic leaflets of the SR are connected to the surface membranes by quasi-periodic bridging structures across a 12–20 nm gap (refs 1,2; see figure). A major unresolved problem is the mechanism by which depolarization (or some other change in the surface membrane) is communicated

across this gap, to increase Ca^{2+} release from the SR lumen into the cytoplasm. Hypotheses to answer this question have been proposed and some, like Peter Pan, retain their youth with sustenance from the Never Land. Will the suggestion that inositol trisphosphate is the messenger across the gap, as put forward for skeletal muscle on page 347 of this issue¹, stay forever young?

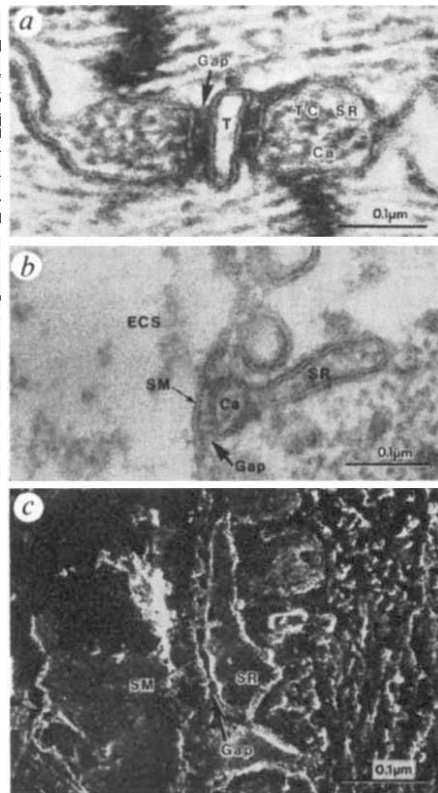
Three main classes of mechanisms have been suggested for excitation-contraction (EC) coupling: first, direct spread of ionic current from the extracellular space through the bridging structures into the SR²; second, mechanical transduction via a mobile protein between the two mem-

brane systems, initiated by voltage-dependent movement of fixed electric charges in the surface membrane⁵; and finally, chemical transmission by a diffusible second messenger. The absence of an ionic current between the extracellular space and the SR together with the fact that Ca^{2+} release is not electrically silent⁶ have excluded the first mechanism^{4,6}, and no experimental test of the second hypothesis has yet been performed. Meanwhile, as the idea of EC coupling by chemical transmission is becoming increasingly attractive, much recent evidence suggests that inositol 1,4,5-trisphosphate (InsP_3) is the second messenger⁷ in many cells. In non-muscle cells, InsP_3 is released by the transmitter-initiated hydrolysis of polyphosphatidyl inositol in the surface membrane and it is thought to mediate transmitter-induced calcium release from the endoplasmic reticulum, a structure that is homologous with SR in muscle. The inevitable question is whether InsP_3 is also the messenger in EC coupling.

When considering EC coupling, it is important to remember that there is an additional mechanism involved in smooth muscle. Whereas in both smooth muscle and striated (skeletal and cardiac) muscle contraction is triggered by electrical depolarization of the surface membrane and its invaginations (the T-tubules) and usually by an action potential, in smooth muscle, neurotransmitters also stimulate contraction or cause relaxation through a 'pharmacomechanical coupling' independent of changes in membrane potential.

Until the discovery of InsP_3 , the most popular candidate as the messenger in EC coupling was Ca^{2+} itself. According to this theory the rapid influx of small quantities of calcium current carried by the action potential triggers the 'calcium-induced calcium release' from the SR, which provides the larger quantity of Ca^{2+} required for initiating contraction. The strongest evidence for calcium-induced calcium release⁸ is in cardiac muscle, much of it from the elegant studies of Fabiato⁹. In skeletal and smooth muscle, however, Ca^{2+} can be released from the SR even in the absence of extracellular Ca^{2+} (and, therefore, Ca^{2+} -influx); nevertheless it is still possible to argue that trigger calcium is released from a protected region, perhaps the cytoplasmic leaflet of the plasma membrane^{4,10}.

The new experiments³ show that in rabbit skeletal muscle, InsP_3 satisfies the first requirement of an EC coupling messenger — it can release Ca^{2+} from the SR. This is demonstrated both with measurements on isolated SR vesicles and by the contractile response of skinned (membrane-free) muscle fibres to InsP_3 . However, such contractions are small, slow and somewhat variable, and seem to require abnormally low free Mg^{2+} concentrations to prevent the rapid breakdown of InsP_3 . None of these objections, as Volpe *et al.* and



Electron micrographs of a triad in striated muscle (a) and surface coupling in smooth muscle (b, c); c has been freeze-etched and rotary shadowed to give a clear view of the junctional gap between the sarcoplasmic reticulum and surface membranes. T, T-tubule, an invagination of the extracellular space in striated muscle; TC SR, terminal cisterna of the sarcoplasmic reticulum; SM, surface membrane of smooth muscle (a from ref. 2; b, c, from ref. 19).

others taking a similar approach¹¹ would argue, are insurmountable. For example, in live muscle, InsP_3 may be released into the diffusionally-restricted junctional space between the SR and surface membranes (see figure); this could cause a rapid and large rise in the local concentration of InsP_3 near the terminal cisternae of the SR where Ca^{2+} is thought to be released. Such regions could even contain lower concentrations of free Mg^{2+} and of the (Mg^{2+} -dependent) phosphatases that break down InsP_3 . If the electron-lucent core of the bridging structures is a lipid, as has been suggested², it could contain InsP_3 precursors. But, for now at least, these are *ad hoc* explanations of the failures of experiments to support the hypothesis.

A more devastating blow to the role of InsP_3 as a physiological messenger of electromechanical coupling in skeletal muscle would be if depolarization were found to increase only the turnover of (mono)phosphatidyl inositol, and not that of the polyphosphatidyl inositol precursor of InsP_3 , and/or if an increase in InsP_3 were shown to be a slow side-effect of acetylcholine released by nerves, and did not precede twitch contraction. The increased phosphatidyl inositol turnover evoked in frog skeletal muscle by depolarization with potassium is far more prolonged than

the duration of Ca^{2+} -release¹², and InsP_3 seems to cause only very modest calcium release from isolated cardiac SR¹³.

The most compelling evidence for a physiological role of InsP_3 in muscle is neither in electromechanical coupling nor in striated muscle, but as a transmitter of pharmacomechanical coupling in smooth muscle. In these tissues various physiological transmitters (cholinergic, adrenergic and peptides) can stimulate the breakdown of polyphosphatidyl inositol to produce InsP_3 (refs 14–16). It is important to note that depolarization does not stimulate breakdown of polyphosphatidyl inositol in smooth muscle^{14,15}, whereas cholinergic agents can stimulate such breakdown during pharmacomechanical coupling¹⁵. The release of calcium from smooth muscle cells by InsP_3 has been demonstrated in two laboratories using different methods^{17,18}. Furthermore, low micromolar concentrations of InsP_3 can cause sustained, graded contractions of smooth muscle even in the presence of the highest free Mg^{2+} (1–2 mM) likely to be present *in vivo*¹⁸. Therefore, it is probable that InsP_3 mediates at least the part of pharmacomechanical coupling that is caused by a voltage-independent release of Ca^{2+} from the SR, and seems to be similar to Ca^{2+} release from the endoplasmic reticulum of non-muscle cells.

It remains to be determined whether the effects of InsP_3 on the SR of skeletal muscle will turn out to be a major physiological mechanism or simply an interesting laboratory phenomenon like the effects of fatty acids and adenine nucleotides, which can release Ca^{2+} or potentiate Ca^{2+} -induced Ca^{2+} -release. The new findings will, at the least, stimulate the hunt for an identifiable messenger of EC coupling. In the background, like the crocodile following Captain Hook, the spectre of a more physiologically active inositol phosphate lurking in InsP_3 preparations haunts investigators in this exploding field. □

1. Franzini-Armstrong, C. & Nunzi, G. *J. Muscle Res. Cell Motil.* **4**, 233 (1983).
2. Somlyo, A. V. *J. Cell Biol.* **80**, 743 (1979).
3. Volpe, P., Salvati, G., Di Virgilio, F. & Pozzan, T. *Nature* **316**, 347 (1985).
4. Hui, C. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2582 (1984).
5. Chandler, W. K. *et al. J. Physiol. Lond.* **254**, 285 (1976).
6. Kitazawa, T. *et al. J. Physiol. Lond.* **350**, 253 (1984).
7. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315 (1984).
8. Endo, M. *et al. in The Regulation of Muscle Contraction: Excitation-Contraction Coupling* (eds Grinnell, A. D. & Brazier, M. A. B.) 181 (Academic, London, 1981).
9. Fabiato, A. *J. gen. Physiol.* **85**, 189 (1985).
10. Bianchi, C. P. *Fed. Proc.* **28**, 1624 (1969).
11. Vergara, J. & Tsien, R. Y. *Biophys. J.* **47**, 351a (1985).
12. Novotny, I. *et al. Physiol. Bohemo slovacca*, **27**, 477 (1978).
13. Hirata, M. *et al. Biochem. J.* **223**, 229 (1984).
14. Akhtar, R. A. & Abdel-Latif, A. A. *Biochem. J.* **224**, 291 (1984).
15. Baron, C. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6899 (1984).
16. Nabika, T. *et al. J. Biol. Chem.* **260**, 4661 (1985).
17. Suematsu, E. *et al. Biochem. biophys. Res. Commun.* **120**, 481 (1984).
18. Somlyo, A. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5231 (1985).
19. Somlyo, A. V. & Franzini-Armstrong, C. *Experientia* **41**, 841 (1985).

Andrew P. Somlyo is Director of the Pennsylvania Muscle Institute, University of Pennsylvania, Philadelphia 19104, USA.

Laser physics

Advances in free-electron lasers

from M.W. Poole

FREE-electron laser (FEL) research is continuing to break new ground; experimental configurations already exist and further modifications are being planned. The potential rewards for the successful development of such a coherent radiation source are very great, whether for radar and telecommunications, fusion energy research, photochemical processing (such as isotope separation) or fundamental research in topics such as non-linear spectroscopy. Ambitious plans are being laid for a 500-Å device and this technology is probably at least as promising as any other for the ultimate goal of an X-ray laser.

The recent announcement that yet another group has joined the increasing band of successful FEL experimenters has significance beyond its initial results. The massive resources of the Lawrence Livermore National Laboratory in California, aided by scientists from Berkeley, have now been brought to bear on the development of this exciting new technology. It seems inevitable that they and their major rivals at the Los Alamos National Laboratory could dominate future progress. Their latest results have demonstrated both exceptionally high gain ($\sim 3,000$) and attractive energy-transfer efficiency (~ 5 per cent) for millimetre waves and the race is now on to extend this to shorter wavelengths.

So far there have been two very distinct FEL communities: those using electron accelerators with beam-current densities sufficiently low that energy transfer is dominated by a single-particle (Compton) interaction, and those with access to intense beams where collective (or Raman) effects become important. The former groups have led the way since the first FEL demonstration by John Madey's team at Stanford University in 1976; important developments were summarized last year in these columns (Pidgeon, C.R. *Nature News and Views* 308, 772; 1984) which noted that four oscillators had operated over the range 0.6–10.6 μm , and since then a fifth device has extended this to 400 μm at the University of California, Santa Barbara. In fact the intense-beam workers could claim a better pedigree that extends back to the ubitron — an early centimetre wave device — and the cyclotron maser that was the centre of attention at the Naval Research Laboratory (NRL) in Washington at the time of the Madey experiments. It was the chance detection of submillimetre output from a cyclotron maser, caused by interaction of the electron beam with its 20-mm generated radiation reflected off a window, that first initiated a long programme of FEL work at

NRL over the past decade, culminating in the reporting of huge gains (120 dB m^{-1}) at 35 GHz (Gold, S.H. *et al. Phys. Rev. Lett.* 52, 1218; 1984). However, all intense-beam FEL experiments until recently have found it necessary to use an additional solenoidal field to focus the large electron beams along the length of the interaction region; such a field can contribute strong cyclotron emission that has led to ambiguities in the interpretation of the output from these devices.

What the Livermore/Berkeley collaboration has been able to announce (Orzechowski, T.J. *et al. Phys. Rev. Lett.* 54, 889; 1985) is a 35 GHz amplifier with high gain and efficiency that exhibits remarkably good agreement with theoretical expectations over a very wide range of signal levels. The experiment was started with the aim of extending collective FEL outputs down to wavelengths similar to those of the Compton devices; this would, for example, allow immediate application of the FEL to plasma heating of fusion mirror machines. Since then, President Reagan's Strategic Defence Initiative has injected financial adrenalin into FEL activities and high power at visible wavelengths is now the declared objective.

The new amplifier, ELF (electron laser facility), is based on the Livermore experimental test accelerator, a linear induction device that can deliver beams of up to 10 kA at 4.5 MeV, although only for short pulses of 30 ns at about 1 Hz. In fact the experiments have been carried out at slightly lower energy and with the electron current heavily collimated to about 500 Å to reduce the beam cross-section in the interaction region. A 3-m-long pulsed electromagnet of 98-mm periodicity induces the required energy transfer between the electrons and an accompanying wave supported by a thin stainless steel waveguide. Additional quadrupoles assist the electron beam focusing but there is no axial field.

First, the team explored the 'super-radiant mode' (laser terminology for amplification from an initial noise signal, in this case the spontaneous synchrotron radiation emitted by a relativistic electron in any magnetic field) and achieved an amplification of 13.4 dB m^{-1} down the periodic magnet. Next, a 60-kW pulse from a magnetron was injected into the interaction region and emerged as a very impressive 80 MW peak output power. In fact the FEL goes into saturation after about 2.2 m and thereafter its power oscillates along the length of the interaction.

A crucial feature of the Livermore approach has been to develop theoretical models in parallel with the experiment-

ation; this has involved extensive computer simulations to include such three-dimensional behaviour as the off-axis electron trajectories and the transverse electromagnetic mode characteristics. Their reward has been an encouraging agreement of predicted and observed FEL behaviour that sets a standard for others to emulate. In the near future the group intends to investigate enhanced efficiency to push the saturated power to still higher levels. At present the efficiency limit is set by the loss of energy from the electron beam detuning the resonant interaction between the charged particles and the wave. To overcome this the Livermore periodic magnet has 15 separate power supplies that can progressively lower the field strength along the interaction region to match the reducing electron energy, thus maintaining the resonance condition. Eventually, Livermore hopes to switch the experiment to an even more powerful induction linac, the advanced test accelerator; this might then permit the same extraordinary performance to be achieved in the infrared region.

Meanwhile, the single-particle FEL community has not been idle. A revolutionary new oscillator in the far infrared has been produced by Elias, one of Madey's original team at Santa Barbara. This takes a giant step towards extremely high average power FEL operation, because the accelerator is a 3 MV Van de Graaff facility that can in principle provide continuous wave electron beams rather than the low-duty cycles typical of most other accelerators. To achieve this would require almost all of the electron beam to be recovered after passing through the FEL, returning it to the high-voltage terminal by a suitable deceleration and collection process; this group has already demonstrated that this can be done. A laser with multikilowatt average power output cannot be far away.

The United Kingdom Heriot Watt-Glasgow-Daresbury collaborative FEL project is now fully assembled at the Kelvin Laboratory, East Kilbride. Spontaneous emission has been detected at both infrared and visible wavelengths, the latter making use of higher harmonic emission generated in the powerful periodic magnet system. Laser oscillation above threshold at 10.6 μm is now being attempted and it is also hoped to demonstrate a wide tuning range. Subsequently to this it is intended to study FEL action on a higher harmonic, a topic of great current interest as it promises to extend FEL operation down to wavelengths previously considered unattainable on such medium-energy (10–100 MeV) accelerators. Apart from this project, sadly almost all funded FEL research is now concentrated in the United States. □

M.W. Poole is in the Synchrotron Radiation Division, SERC Daresbury Laboratory, Warrington WA4 4AD, UK.

Pesticides and the decline of Guam's native birds

SIR—Diamond¹ recently reviewed the factors which might be responsible for the decline of Guam's native birds. Although Diamond's commentary was based on a report by Jenkins² that included several hypotheses (pesticides, introduced predators and disease), Diamond suggested that pesticides were the most likely cause.

The pesticide hypothesis evolved from two reports. In 1946, Baker³ stated the organochlorine pesticide DDT was used extensively on the island during and after the Second World War. In 1977, Drahos⁴ reported concentrations of DDE, a metabolite of DDT, of less than 0.4 p.p.m. in the carcass lipid (about 0.03 p.p.m. whole carcass wet weight) or swiftlets (*Aerodramus vanikorensis bartschi*) and concentrations of DDT and DDE of less than 0.10 p.p.m. (believed to be on a dry weight basis) in swiftlet guano.

Diamond¹ and others^{2,4} considered these data indicative of potential harmful effects of DDT, a conclusion not supported by critical evaluation of Drahos' data or concentrations of DDT and DDE (≤ 0.2 p.p.m. dry weight) detected in swiftlet guano I collected in 1981. The carcass residues reported by Drahos⁴ indicate that swiftlets were exposed to DDT or DDE, but concentrations are a fraction of those associated with mortality (> 600 p.p.m. in carcass lipid)⁵ or reproductive failure (78 p.p.m. whole carcass wet weight)⁶ in birds. Residues in the guano are difficult to interpret because relationships between concentrations of organochlorines in bird faeces and those in their carcasses have not been determined. These relationships have been determined for bats, however, and residues in swiftlet guano are similar to those in guano of bats (*Myotis grisescens*) whose carcasses contained only a fraction of lethal levels⁷.

Soil and avian prey I collected in 1981 were essentially free of organochlorine pesticides; only a sample of skinks (*Carlia fusca*) contained a detectable concentration (DDE at 0.04 p.p.m. wet weight, 0.13 p.p.m. dry weight). Residues in shrews (*Suncus murinus*) trapped in 1981 did not exceed 0.30 p.p.m. wet weight. These levels are also less than those known to adversely affect birds or small animals.

In addition, it is unlikely that present use of pesticides by civilians or the military on Guam is responsible for the decline. Agricultural lands comprise less than 1% of the island's total area and more than 84% of the pesticide used is organophosphates or carbamates⁸, chemicals that do not persist in the environment⁹. Of the four pesticides used the most (malathion $>$ diazinon $>$ naled $>$ carbaryl), only diazinon has been implicated in wildlife mortality elsewhere⁹. Malathion is also the pesticide used in the greatest quantity by the military, and with few exceptions,

use has been restricted to housing areas. Although spray operations by the military may have temporarily decreased insect abundance, such decreases cannot explain the continuing decline of the island's avifauna.

It is unfortunate that Diamond¹ and others^{2,4} incorrectly equated pesticide use or residues on the island with harmful effects. Although the use of some pesticides can be hazardous to birds, and effects associated with the use of DDT still persist today^{9,10}, there are no data supporting the hypothesis that pesticides are responsible for the continuing decline of Guam's native avifauna.

CHRISTIAN E. GRUE

US Department of Interior,
Fish and Wildlife Service,
Patuxent Wildlife Research Center,
Laurel, Maryland 20708, USA

1. Diamond, J. M. *Nature* **310**, 452 (1984).
2. Jenkins, J. M. *Orn. Monogr.* No. 31 (1983).
3. Baker, R. H. *Trans. N. Am. Wildl. Conf.* **11**, 205-213 (1946).
4. Drahos, N. *The Status of the Vanikoro Swiftlet on Guam* (Guam Division of Aquatic and Wildlife Resources, 1977).
5. Stickle, L. F. & Stickle, W. H. in *Pesticides Symposia* (ed. Diechman, W. B.) (Helios, Miami, 1970).
6. Mendenhall, V. M., Klaas, E. E., McLane, M. A. R. *Archs Envir. Contam. Tox.* **12**, 235-240 (1983).
7. Clark, D. R. Jr, LaVal, R. K. & Tuttle, M. D. *Bull. Envir. Contam. Tox.* **29**, 210-220 (1982).
8. *The Impact of Agricultural Activity on Guam's Water Quality* (Guam Environmental Protection Agency, 1983).
9. Grue, C. E., Fleming, W. J., Busby, D. G. & Hill, E. F. *Trans. N. Am. Wildl. nat. Resour. Conf.* **48**, 200-220 (1983).
10. Fleming, W. J., Clark, D. R. Jr & Henny, C. J. *Trans. N. Am. Wildl. nat. Resour. Conf.* **48**, 186-199 (1983).

SIR—In his *News and Views* piece on the "Possible effect of unrestricted pesticide use on tropical birds"¹, Diamond used Guam's avifauna as an example based on Jenkins' report *The Native Forest Birds of Guam*². Jenkins suggested that pesticide usage may be responsible for the drastic decline of Guam's birds. However, due to publication time lag, Jenkins' monograph is only current through 1979. Research since that time suggests that although pesticides were extensively used on Guam, they do not appear to be responsible for the general decline of the avifauna (C. Grue, personal communication). However, pesticides may have impacted certain species in the past, such as the insectivorous Vanikoro swiftlet and sheath-tailed bat.

My research has focused on the possible roles of introduced avian diseases and predators. Between 1982 and 1984, a variety of birds were sampled in collaboration with the USFWS National Wildlife Health Laboratory for parasites, bacteria and viruses and two sentinel experiments were conducted. To date, no infectious organisms have been isolated that could account for the current extinctions and range reductions.

Feral dogs, cats and rats are a problem on all major islands in the Marianas. The only predator unique to Guam is the brown tree snake, *Boiga irregularis*. The snake is the major cause of the decline of

the forest avifauna. This species was introduced to Guam in 1946, and its range expansion corresponds to the avian range contraction. Predation intensity on bird-baited traps is extremely high, and alternative prey items have also declined. Supporting data are currently being prepared for publication.

Thus, although Guam may not be the best example for the dangers of unrestricted pesticide usage, it certainly should serve as a warning of the potential disastrous effects of introducing exotics to island ecosystems.

J. SAVIDGE

Division of Aquatic & Wildlife Resources,
Guam, Marshall Islands 96913, USA
and
Ecology, Ethology and Evolution,
University of Illinois,
Champaign, Illinois 61820, USA

1. Diamond, J. *Nature* **310**, 452 (1984).
2. Jenkins, T. M. *Am. Orn. Un. Orn. Monogr.* No. 31 (1983).

Electrostatics revindicated . . . classically . . .

SIR — Berezin¹ has considered the minimum energy configuration of N point charges placed inside or on a circle. This problem has a well defined continuum limit, namely the distribution of charge over a metallic disk. The solution, given in Jeans' textbook², has a finite charge density all over the disk rising to infinity at the edge. Consistency with this limit demands that for increasing N , charges (in the minimum energy state) leave the edge and occupy the interior. The use of a $1/r$ rather than an $\ln r$ potential implies that the circle is really a disk embedded in three-dimensional space and there is thus no contribution with the usual notion that charges reside on the surface of a conductor. These arguments imply that in the case of a sphere, there will be no minimum energy solution with charges in the interior, as is already clear from Earnshaw's theorem³.

R. NITYANANDA

Raman Research Institute,
Bangalore-560 080, India

1. Berezin, A. A. *Nature* **315**, 104 (1985).
2. Jeans, J. *The Mathematical Theory of Electricity and Magnetism*, 167, 249 (Cambridge University Press).

. . . by conducting ellipsoid

SIR — A very old problem in electrostatics is to find the distribution of charge on a conducting ellipsoid which will be taken to be an ellipsoid of revolution. The charge density is of course confined to the surface and is everywhere finite. If we let the length of the axis of rotation go to zero we get a highly idealized two-dimensional object, a two-dimensional (infinitely thin) conducting disk. If the radius of this disk is R , the charge density is well known to be proportional to $(R^2 - r^2)^{-1/2}$. This density

must be a fairly good approximation to the distribution of a number of identical charges (say electrons) placed on the disk. It is not confined to the circumference and it is *not* zero at the centre. Hence there is really nothing unexpected in the results given by Berezin¹ who shows that 12 identical charges distributed evenly on a circle have a greater potential energy than 11 charges on the circle and one at the centre. Perhaps the fact that 12 is the precise number at which one charge is "expelled to the centre" is mildly surprising.

Berezin goes on to conjecture that a similar "spontaneous ejection" of charge to the centre will occur in three dimensions and that this "may lead to some modification of usual theorems of electrostatic stability which claim that at the state of equilibrium all charges . . . are always located at the surface." I think that this conjecture is wrong, and is a nice example of incorrect generalization from two to three dimensions. For example, arguing from two² to three dimensions about the propagation of an impulsive wave emitted from the origin would lead one to the false conclusion that there is a residual disturbance left at the centre in the three-dimensional case.

However, Berezin's letter raises a question to which I have been seeking an answer for some time. It seems clear that it is possible to place an arbitrary number, N , of electrons on a conducting sphere. The question is how do they arrange themselves?

There is little doubt in my mind that they will arrange themselves on the surface, and intuitively the answer would seem to be given by simple symmetry arguments when $N=1, 2, 3, 5$ and $4, 6, 8, 10, 20$, where the second set of numbers are the numbers of vertices of the five regular polyhedra in three dimensions. However, as Berezin has pointed out (in personal communication), such arguments are false in the case where $N=8$. Having the electrons at the vertices of a regular cube has a *higher* energy than when one face is rotated relative to the opposite face by 45° , the distance between the faces being held constant. A quantum treatment of this problem would also be interesting and could also yield mean values of the multipole moments of the system.

A. M. CORMACK

Physics Department,
Tufts University,
Medford, Massachusetts 02155, USA

1. Berezin, A. A. *Nature* **315**, 104 (1985).
2. Morse, P. M. & Feshbach, H. *Methods of Theoretical Physics*, 1363 (McGraw-Hill, New York, 1953).

... with charged polygons

SIR—The classical electrostatic energy of N point charges, q , symmetrically disposed on a circular circumference with radius R has been recently discussed here (A. A. Berezin, *Nature* **315**, 104; 1985). A surprising feature of these regular poly-

gonal systems was reported for $12 \leq N \leq 400$. Such arrangements are energetically unstable with respect to displacement of one charge to the centre and regular circumferential redistribution of the remaining $(N-1)$ charges. On this basis Berezin made two conjectures: (1) that such a rearrangement is energetically favourable for any large N and (2) that his large- N result indicates a violation of the familiar electrostatic theorem. For a conducting body at electrostatic equilibrium, any net charge must be located at the surface (Faraday shielding).

It should be pointed out that conjecture (2) cannot follow from the large- N polygon results.

If these systems are to model a classical conducting system, their electrostatic potential energy should vary as: $W = \frac{1}{2} cq^2 N(N-1)$ with c a geometrically determined constant (capacitance coefficient). This in turn requires that the electrostatic potential at any vertex location due to the other $(N-1)$ charges should be $U = cq(N-1)$. Evaluation of exact electrostatic potentials for polygons with $N=10^3, 10^4, 10^5$ yields respectively $U=2,238.7970; 29,717.3261$ and $370,466.8178$ (units of q/R).

Clearly the required linear N dependence does not occur. This failure relates to the usual logarithmic distance dependence of electrostatic potentials near linear charge distributions. Thus the polygonal charge arrangements cannot model a conducting system.

For the problem at hand, the potential U may be approximated by exactly summing the contributions of the $2j$ charges closest to a vertex location and treating the remaining charges continuously. For example, choosing $j=18$ and N large:

$$U \approx \frac{Nq}{R} \left[1.112527455 + \frac{1}{\pi} \ln \cot \frac{37\pi}{4N} \right]$$

This expression reproduces the potential energy, U , for polygons with $N \geq 10^3$ with fractional accuracy better than 2.3×10^{-5} ; this accuracy improves with increasing N . The non-linear N dependence is manifest. Using such an algebraic approximation Berezin's first conjecture, the instability of any large- N polygon system, is readily confirmed.

For N surface charges regularly distributed on the surface of a sphere with radius R , the limiting form of the electrostatic energy is:

$$W = \frac{1}{2} \frac{q^2}{R} (N) (N-1)$$

Accordingly for the spherical surface, expulsion of one charge to the centre is energetically possible only in the infinite N limit.

R. A. NAUMANN

Physik-Department, E18,
Technische Universität München,
D-8046 Garching b. München, FRG

... but for 12 equal point charges?

SIR—Berezin has recently shown that the minimum energy configuration for N equal point charges placed in a circle is different depending on whether N is less than, greater than or equal to 12 (*Nature* **315**, 104; 1985). In the former case ($N < 12$) the configuration of minimum energy is with the N charges at the vertices of a regular polygon inscribed in the circle (configuration A); in the latter case ($N \geq 12$) a lower energy configuration is obtained with one charge expelled to the centre of the circle and $(N-1)$ charges at the vertices of an $(N-1)$ sided inscribed polygon (configuration B). Berezin verified this result up to $N=400$ and claimed it was "very likely" true for all $N \geq 12$. It is easy to show the result does indeed hold for all $N \geq 12$ as follows.

The energy of N charges arranged in configuration A is

$$W_A(N) = \frac{N}{2} \left\{ \sum_{i=1}^{(N-1)/2} \frac{1}{\sin \pi i/N} + 0.5 (N \text{ even}) \right\} \quad (1)$$

and arranged in configuration B is

$$W_B(N) = \frac{(N-1)}{2} \left\{ \sum_{i=1}^{(N-2)/2} \frac{1}{\sin \pi i/(N-1)} + 2 + 0.5 (N \text{ odd}) \right\} \quad (2)$$

It follows simply from equations (1) and (2) that

$$W_B(N) = W_A(N-1) + (N-1) \quad (3)$$

Hence the condition that configuration B gives higher energy than configuration A ($W_B(N) > W_A(N)$) is true for those N where

$$W_A(N) - W_A(N-1) < (N-1) \quad (4)$$

By inspecting the numerical data for $W_A(N)$ only, looking say at the range $2 < N < 40$, it is possible to see that the inequality (4) is satisfied for $N < 12$ as shown by Berezin. It is also easy to see that the left hand side of inequality (4) is increasing faster than $(N-1)$ as N runs beyond 12. Hence the result found by Berezin is true for all $N \geq 12$.

STEVE WEBB

Joint Department of Physics,
Royal Marsden Hospital,
Downs Road, Sutton,
Surrey SM2 5PT, UK

The rise of 'Mobilism'

Roy Porter

The Dark Side of the Earth. The Battle for the Earth Sciences: 1800-1980.

By Robert Muir Wood

George Allen & Unwin: 1985. Pp.246. £11.95, \$19.95.

"GEOLOGY is dead; long live the Earth sciences". This in a nutshell is the message of Robert Muir Wood's analysis of continental drift from Wegener's chequered career up to the sudden and stunning success of "mobilism" in the plate tectonics revolution of the 1960s.

At one level, this bold survey can be read with equal pleasure and profit as one of the best interpretations of that astonishing scientific upheaval. Unlike Anthony Hallam's *A Revolution in the Earth Sciences* (Oxford University Press, 1973) it does not primarily aim to evaluate the inner logic of successive geological theories; nor does it rival the dramatic blow-by-blow commentary on tectonic breakthroughs offered by William Glen's *The Road to Jaramillo* (Stamford University Press, 1982). But Dr Wood's reading has special virtues of its own. He offers the best account yet of the nineteenth century heritage of speculations about oceans and continents which formed the matrix of drift theory. And he is adroit in bringing his protagonists to life and illuminating their particular parts in the story. Harry Hess is identified as the senior figure who consistently championed younger researchers such as Fred Vine in their heterodox investigations; John Tuzo Wilson is shown to have been the right man in the right place at the right time for propounding the wider "mobilist" synthesis. And due weight is given to the crucial role played by Teddy Bullard as a catalyst, above all in making Anglo-American co-operation in this enterprise so fruitful. Not least, Dr Wood commands a good turn of phrase, and patiently weaves together with enviable skill all the strands of a highly complex narrative.

So here is a sure-footed and wide-ranging account of the triumph of continental drift, with much to offer to specialists and non-specialists alike. But Dr Wood has bigger fish to fry. His case is that the plate tectonics revolution was not just a great transformation in geology — indeed, it is wrongly interpreted if we see it as a revolution in geology. For it was rather a revolution *against* geology.

From its early nineteenth century giants up towards the present, geology's great triumph lay in the stratigraphical and palaeontological vision, in generating the science of the rocks, indeed, a love of the rocks. Geology studied the land masses, especially their mountains. As synthesized in classics such as Eduard Suess's *Das Antlitz der Erde* (1885-1909), it theorized the continents in the context of a

stable, if shrinking, Earth, and then minutely studied their deformations through tireless fieldwork and mapping.

The plate tectonics revolution has left this either wrong or stranded high and dry as an irrelevance. The "mobilist" vision swung attention away from the establishment's preoccupation with hammering out ever more data about the strata. Instead "drifters" turned the whole Earth into their parish and their problem. Unlike conventional geologists, they followed Lyell's advice (think like intelligent amphibians) and gave the ocean beds at least as much attention as the land masses. They insisted on tackling questions, such as the global distribution of the continents, from which orthodox geology shied away; they deployed specialist skills in fields such as seismology and geomagnetics not routinely possessed by pukka geologists; and they ended up proposing a theory which could not be judged — for or against — at the bar of conventional geology.

In other words, the revolution in plate tectonics was the overtaking of geology by an alliance of disciplines which we may call geophysics. Or, put another way, enter the Earth sciences, marching under the banner of a global theory, which rendered old geology obsolescent in much the same way as Darwin's evolutionism left behind classificatory natural history as a key to the economy of life.

This was a revolution made by aliens. Some were geographical outsiders, such as the South African, Alexander Du Toit, or Warren Carey from Australia. Others were disciplinary marginal men, classically of course Wegener himself (first and foremost a meteorologist, as Dr Wood rightly insists, while perhaps unduly playing down his geological credentials). Practically all the pioneers of "mobilism" had training or research experience in some branch of geophysics — often in seismology or geomagnetism — which set them apart from true-blue geologists. Not least, "mobilism" was typically the brainchild of the "mobile", the young and ambitious, and was perceived by the old guard to be unsettling (in America it was dubbed "leftist", in Russia, "bourgeois").

Thus the plate tectonics revolution did not emerge out of the geological main stream. Above all, its inspiration came from that late arrival, deep-sea geology. And this in itself would have been unthinkable without the gigantic sea-bed surveying operations funded at huge expense by allied governments during and

after the Second World War (projects in which Hess, Ewing and Bullard were all prominent); and then, at a later stage, without the United States pouring money into seismology to monitor nuclear tests. Thus strategic needs produced new fields of expertise; these led to the discovery of the magnetic field anomalies of seafloor rocks, which in turn proved the surprising youth of the ocean floor, and gave rise to the key idea of seafloor spreading, that *since qua non* of the vision of continents borne on mobile plates.

Of course, Dr Wood is not arguing that conventional geology contributed nothing at all to this. But the fact that even so catholic and prescient a geologist as Arthur Holmes got so far yet no further in convincing himself and others of drift, perfectly shows its limitations even at its best. And, at its worst, orthodox geology proved to be an ostrich.

So plate tectonics should not be seen as a revolution in geology, but as marking geology's eclipse, a refocusing of scientific attention away from the rocks towards the whole Earth, a re-think of planet Earth for the space age. Science thus advances, Dr Wood argues, not by accumulating data, nor even by internal theory switches, but by massive re-groupings of disciplines. The old gods are devoured not by their own children but by their neighbours' (except perhaps in the Soviet Union where geological gerontocrats have had more success in keeping the new "drifters" at bay). Late Victorian geology had resented being put in its place by the physico-mathematical imperialism of Lord Kelvin, and had unilaterally declared its independence of physics. Now, a century later, geophysics has had its revenge.

Occasionally, the contrast Dr Wood draws between senile geology and virile Earth science is overstated for polemical effect (his own evidence sometimes gives it the lie); and it is a pity that, drawing too freely upon hindsight, he oversimplifies certain complex figures, such as T.C. Chamberlin and Maurice Ewing, and turns them into whipping boys. Although providing attractive line drawings of the main protagonists, the book would have benefitted, I think, from some explanatory diagrams. Yet all praise to him for recognizing that in understanding the history of science no less than in understanding the history of the Earth, "fixism" will not do. No more than the continents, are the sciences themselves fixtures; they too are subject to drift, expansion and subduction. This is an important re-interpretation, whose "mobilist" vision should send some shock waves through the geological establishment. □

Roy Porter is Senior Lecturer at the Wellcome Institute for the History of Medicine, Euston Road, London NW1 2BP, UK and author of *The Making of Geology* (Cambridge University Press 1977).

Getting down to the bare bones

Richard Potts

The Analysis of Animal Bones from Archeological Sites. By Richard G. Klein and Kathryn Cruz-Urbe. *University of Chicago Press: 1984. Pp.266. Hbk \$19.95, £18.95; pbk \$8.95, £8.50.*

INVESTIGATION of past human activities and ecological settings requires many analytical steps and cautious weighing of the processes that transform living phenomena into buried, fossilized traces. Archaeological faunal remains—one such “end product” of this transformation—have provided extremely important, albeit controversial, evidence about early hominid behaviour and environments. After determining the species, body part, sex and age of death of the animals represented by bone specimens, zooarchaeologists often attempt to estimate the relative number of individuals or bone units of each species found in an accumulation of bones; further, they may construct mortality/age distributions for well-represented species. These and other quantitative data help to test ideas about environmental change; catastrophic versus attritional death of animals; whether humans, carnivores or geological agents were responsible for bone aggregates; and whether humans hunted or scavenged, and how they selected certain species and body parts.

The theme of this volume is the entire process by which zooarchaeologists interpret the past from the assemblages of vertebrate bones often found in

archaeological excavations. The first of two sections introduces lucidly the fundamentals of animal bone identification, faunal data collection and analysis. The second provides computer programs in BASIC that generate data used widely in faunal analysis (minimum number of individuals, age distributions and so on). These programs are explained clearly and the procedures used to quantify bone assemblage characteristics are discussed thoroughly. Inferences about ancient people, their sites and environments are illustrated largely from Klein's own research on prehistoric faunas of southern Africa and Europe. Thus, the book provides an excellent summary of the approaches Klein has developed.

In illustrating certain aspects of faunal analysis, the authors offer interpretations that appear *post hoc*, without due consideration of alternatives; this will disquiet some researchers. Others may feel that the authors' distinction between the comparative and taphonomic approaches to faunal studies is unwarranted, especially since the latter approach (which examines how modern processes create faunal patterns) is played down.

These potential points of dispute do not overshadow the book's value, however. It will be widely used as a text on fundamental aspects of faunal analysis and as a source of ideas and useful programs for generating data that help to explore the subtle, bony traces of ancient human activities and environments preserved in the geological record. □

Richard Potts is Associate Curator in the Department of Anthropology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA.

The wonderful sky at night

Martin Harwit

The Cambridge Atlas of Astronomy. Edited by Jean Audouze and Guy Israël. *Cambridge University Press: 1985. Pp.432. £29.95, \$75.*

How wonderful to open this volume to find an absolutely lavish collection of new drawings, charts, explanatory diagrams, illustrated tables and photographs—not just of astronomical objects, but also of terrestrial fossils and atmospheric and volcanic structures.

There are charts and maps of all kinds, and pictures of the lunar surface and of all the planets and satellites visited by spacecraft. Some of these we have seen before, but most of them are selected to be instructive: tracks left on the Moon's surface by the Apollo lunar vehicle, along-

side lunar dust, highly magnified, showing an occasional spherical droplet, and next to that an electron micrograph of a single grain showing tracks of solar cosmic rays that have bombarded the lunar regolith leaving a record of solar activity over past aeons. Over the page we find a picture of a micrometeorite, followed by views of slices of larger meteorites and embedded ion tracks. Just by looking at such sequences readers will gain insight into the nature of lunar surface material, meteoritic matter and some of the processes to which both have been subjected. In turn, they will better appreciate the way a knowledge of these processes helps us to understand the past histories of the materials which make up the Solar System.

This, then, is a beautifully produced book of great pedagogic value. Amateur astronomers, high school students and professionals will find here a visual display which is both arresting and informative. What I am about to say should, therefore, not discourage one from investing in this volume. □

My particular disappointment started when I stopped looking at the illustrations and began reading the text. In just the first hour of browsing I noted a dozen flaws of all kinds. There are numerous typographical errors, and the index appears to have been compiled with insufficient care over the cross-referencing, which is so important in a volume such as this. More specifically, a figure caption (p.11) describes the telescopes used by Galileo, and adds, “The principles of these telescopes were conceived in Venice by an unknown genius”. Actually, these instruments first surfaced in the Netherlands, late in 1608, roughly a year before Galileo began his observations. And on p.387 we find the statement: “The existence of neutrinos was proposed for the first time by Paul Dirac on a theoretical basis to explain the conservation of the angular momentum, or spin of the ensemble of particles produced by the beta disintegration of the neutron”. Actually it was Wolfgang Pauli, with some help from Enrico Fermi in naming the particle, who conceived of the neutrino for conserving energy in beta decay.

More serious, however, is the question of purpose. How much should the text explain—couldn't the occasional bit of calculus have been avoided through use of a simple diagram? One has the impression that the editors wanted to pack every possible concept into this single volume—massive neutrinos, the number of neutrino generations, proton decay and other current exotica. This has taken its toll on clarity. For example the section on cosmology begins by listing five principles alleged to guide the construction of cosmological models: “The principle of uniformity, the cosmological principle, the anthropic principle, the principle of equivalence and Mach's principle” (p.381). But Mach's principle still does not fit into any of our current theoretical models; and the anthropic principle merely tells us to disregard theoretical models which don't ultimately produce cosmologists. Everything seems to have been crammed in with little regard for separating the important from the frivolous.

Still, the editors and publisher have done remarkably well with the illustrations and, on balance, I recommend the *Atlas* highly. I just hope that the text can be thought through once more, in time for a well-deserved second edition. □

Martin Harwit is a Professor in the Department of Astronomy, Cornell University, Ithaca, New York 14850, USA.

New in paperback

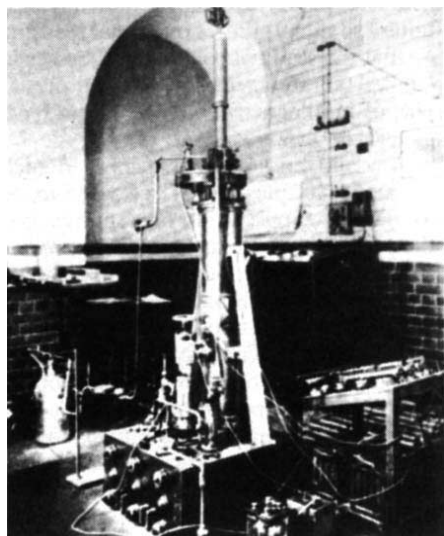
- *The Dawn of Animal Life: A Biohistorical Study* by Martin F. Glaessner. Cambridge University Press. £9.95, \$19.95.
- *Introduction to Quantum Mechanics with Applications to Chemistry* by Linus Pauling and E. Bright Wilson, Jr. Dover. \$9.95.

The unseen revealed

A. Howie

The Beginnings of Electron Microscopy. Edited by Peter W. Hawkes. *Academic*: 1985. Pp.633. \$88, £67.

DURING the 1920s, the work of Busch and Gabor on the focusing action of short magnetic solenoids in the high speed cathode ray oscillograph, and the separate, celebrated discoveries in electron diffraction, sparked off an idea which some minds — with a faith reminiscent of a mediaeval cathedral-builder — took seriously. That idea was the construction



The Met-Vick EMI, the first commercially produced electron microscope. Installed at Imperial College, London in 1936.

of microscopes to surpass the resolution of optical instruments by as much as five orders of magnitude, a limit deduced purely from the electron wavelength.

Here, in partially coherent illumination from a variety of authoritative sources (delightfully illuminated themselves by Cosslett), one can learn how the idea survived the technical obstacles of power supply instability and poor vacua, as well as the deeper problem of lens aberrations. It seems that other difficulties, such as the slump of the early 1930s and the Second World War (with enemy occupation of Holland and France to be borne by Le Poole, Grivet and Dupouy), were only marginally more serious than the general indifference of optical microscopists. Did they simply exaggerate the beam damage problem? Or were they more realistic about the looming difficulties of specimen preparation and image interpretation described here by Drummond, Hall and Fernandez-Moran?

In this enjoyable collection of reminiscences, Dr Hawkes has unfortunately missed some potentially important contributions because of illness, death or prior publication elsewhere. The brilliant construction by Prebus and Hillier, two

graduate students at Toronto, during four months in 1938, of an instrument eventually achieving 100Å resolution is only briefly mentioned. Less serious is the absence of a major contribution from Ernst Ruska (whose own book, *The Early Development of Electron Lenses and Electron Microscopy*, appeared in English translation five years ago). His towering achievement is manifest here from different viewpoints. As noted by Mulvey, the transmission microscope producing today's images at atomic resolution follows in all crucial respects the original Ruska design. Restricted, however, by lens aberrations to the small-angle range where coherent scattering effects dominate, the instrument depends on a marriage of electron optics and electron diffraction far closer than anyone then envisaged. Although the reader may trace in Le Poole's account the instrumental foundations of that partnership (also laid by Boersch), sample its fruits in Dupouy's magnificent high-voltage images or sense with Hibi the holographic future, he must turn elsewhere (for example to *Fifty Years of Electron Diffraction*, published by Reidel in 1981) to appreciate the enormous benefits which the union brought to both disciplines.

As its originator von Ardenne describes, scanning electron microscopy provides a contrapuntal theme even though it shares the same origins with the more conventional transmission method. A much better understanding of electron beam scattering processes is required, however, to exploit fully the rich range of imaging modes available in scanning. Its wide application depends more on ease of specimen preparation and depth of field than simply on resolution. The post-War exploration of these questions and the development of instruments for manufacture are recounted by Oatley and his colleagues. This contribution, as well as Mulvey's, covers two of the three principal manufacturing initiatives in electron optics where, contrary to received doctrine, British industry exploited an idea originating elsewhere. Le Poole and Kanaya treat the early stages of the important Dutch and Japanese electron optical industries.

A third theme, dating from pre-War Germany but little mentioned here, is the field emission microscope. Closer to the simplicity of the Leeuwenhoek microscopes than anything else in electron optics, the field emission tip has fathered the recent spectacular advances in scanning transmission and in scanning tunnelling microscopy. Before long, perhaps, Dr Hawkes will edit a second volume rediscovering in this more recent progress our great debt to the persistent vision of the early pioneers. □

A. Howie is Reader in the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, and President of the Royal Microscopical Society.

Plumbing the depths

Stephen Moorbath

Archaean Geochemistry: The Origin and Evolution of the Archaean Continental Crust. Edited by A. Kröner, G. N. Hanson and A. M. Goodwin. *Springer-Verlag*: 1984. Pp.286. DM87.

UNTIL recently, the Archaean (c. 3,800–2,500 Myr ago) was still a great unknown in the earth sciences. This situation has changed dramatically since the 1960s as a result of two main factors: (i) advances in isotope dating methods and their application to Precambrian rocks, and (ii) the recognition that the oldest exposed gneiss and greenstone terrains retain a wealth of primary igneous, sedimentary and geochemical features.

It was clearly difficult to find an accurate title for a book which contains 13 miscellaneous contributions to a wide field, ranging in style from reviews to research papers and arguably more suited for publication in journals. Thus the principles of geochemistry were no different during the Archaean than they are today. Some workers (including myself) even consider that Archaean continental crust originated and evolved by processes closely similar to those of geologically recent times, albeit much more rapidly because of greater terrestrial radiogenic heat

The Ecology of Rocky Coasts

Edited by Dr P G Moore and Dr R Seed

This is probably the most comprehensive volume on the subject of the ecology of rocky coasts ever published. The book is the work of 28 internationally outstanding experts in the field and includes the most recent research and evaluations in this increasingly important area. It provides a complete overview of littoral and sub-littoral communities which will be of immense value to ecologists and marine biologists.

0 340 37011 4 Illus. Boards
£40.00

Available through all good bookshops

Hodder & Stoughton

Mill Road, Dunton Green, Sevenoaks, Kent
TN13 2YD

Reader Service No.7

production. The deep crustal level represented by many Archaean rocks now exposed at the surface conveniently gives us insight into the chemical and petrogenetic processes occurring in the deeper reaches of the continental crust, not normally observable at those accreting or colliding plate margins where young continental crust is nowadays being created or re-worked.

Certain aspects of the subject are efficiently dealt with in this useful volume, whose stated aim is to summarize some of the main issues in the broad geochemistry of Archaean rocks which were addressed during a nine-year stint of one of the projects ("Archaean Geochemistry") of the International Geological Correlation Programme. Topics range from the accretion history of the Earth and composition of the primordial mantle to greenstone belt evolution, petrogenesis of komatiites, the nature of high-grade gneiss terrains and the role of fluids in their development and geochronology, and the relation of Archaean sediments to the overall composition of the Archaean continental crust. Also included are geochemical and geochronological data on regions in the Soviet Union, China and India which were hitherto little known in the Western literature. It is refreshing, for once, to see some reliably documented new isotopic age data from the oldest shield areas of the Soviet Union to which up-to-date analytical and interpretative techniques have been applied.

I regret the absence of a contribution on the rapidly developing field of biogeochemistry and related topics, since the Archaean saw the transformation of "primordial soup" into the highly organized, far-reaching second course. Some discussion of Archaean metallogenesis would also have interested geochemists. □

Stephen Moorbath is Reader in Geology in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Red Tam nestling in the Helvellyn corrie. From Discovering Landscapes in England and Wales by Andrew Goudie and Rita Gardner. Published by George Allen & Unwin at £12.95.

Animal cognition

Geoffrey Hall

Animal Intelligence. Edited by L. Weiskrantz. Clarendon: 1985. Pp. 223. £32.50.

THE publication of this work (the proceedings of a Royal Society discussion meeting) just over a century after the appearance of *Animal Intelligence* by G. J. Romanes tempts one to make comparisons. Romanes was interested in animal intelligence because he felt the findings of comparative psychology had as much relevance to Darwinian theory as did those of comparative anatomy.

Romanes would be at home with Jerison's contribution to the present volume, a chapter which charts the evolution of encephalization and the implications of brain size for mental activity. But, as Jerison points out, little use is made elsewhere of evolutionary theory in organizing the wide range of facts about animal behaviour presented by the various contributors. One of the reasons for this neglect may be discerned in the chapters by Macphail and by Mackintosh and his collaborators. Although Romanes felt able to rank the various animal groups according to their level of intelligence (his scheme is reproduced in Weiskrantz's introductory chapter), modern experimental studies of animal learning (such as are reported by Mackintosh) have found precious few differences between species that can be unambiguously interpreted as being the result of differences in intelligence. Macphail concludes that we might as well accept what he calls the "null hypothesis" — that extant non-human vertebrates do not in fact differ in intelligence.

Romanes has been castigated by subsequent comparative psychologists for basing his conclusions on mere anecdote rather than on careful observation and experiment. To some, Macphail's conclusion may seem not so much an accurate judgement of the intelligence of animals as a condemnation of the skills of modern experimenters. And certainly, some of the most striking examples of intellectual achievement among animals reported in this book come not from laboratory studies but from observations made in natural or semi-natural conditions. Goodall concludes her intriguing survey of the innovative behaviour displayed by the chimpanzees of Gombe with an assertion of the value of (carefully documented) "anecdotal" reports of natural behaviour.

However this may be, none would deny the need for experimental analysis in any attempt to determine the mechanisms responsible for seemingly intelligent behaviour and much of this book is devoted to reports of such work. Several chapters present accounts of the behaviour shown by primates when solving the problems set for them by experimental psychologists. It

is made clear that their behaviour cannot be interpreted simply as the acquisition of new habits and must require intelligence of a higher order. But perhaps this has never really been in doubt. More significant, therefore, is the elegant analysis presented by Dickinson of the mechanisms by which that least esteemed of animals, the laboratory rat, learns the simplest of tasks, that of pressing a lever to obtain food. What emerges is a need to allow the rat knowledge about the relationship between its actions and its desired goals, and credit it with the ability to carry out inferences on the basis of that knowledge. No more than this is claimed for the primates, a fact that not only speaks in favour of Macphail's "null hypothesis" but also helps to justify the belief that has sustained so many studies of animal learning — that the study of cognitive processes in one vertebrate species will reveal principles of intellectual functioning that have general applicability.

It would be a disappointment if such general principles as emerge were to be found not to apply to our own species. But the possibility remains that human intelligence is different in kind from that shown by non-human animals, being based on our possession of a form of language that is beyond the capacity of any beast. Empirical investigation of this hoary notion long seemed impossible. But perhaps the most interesting implication of the recent attempts by several experimenters (Terrace, Gardner and Gardner, Savage-Rumbaugh and others are represented in this book) to teach a form of human language to an ape is that their work appeared to allow just such a possibility. Unfortunately, the outcome has been inconclusive. There is no real consensus as to what the apes have actually achieved or, more worrying, as to what it is that they would have to do in order for us to be sure that they had learned a language. An empirical solution seems as far away as ever.

More generally, this book leaves one with the impression of a field of study that is productively active, has achieved much in the hundred years since Romanes and is likely to achieve more in the near future. And this is only half the story. Romanes devoted roughly half of his book to the intellectual achievements of invertebrates. That these animals receive no mention in the present volume is its only serious failing. It is now becoming clear that the application of methods of study and analysis previously applied only to vertebrates can reveal evidence of equivalent intellectual sophistication in invertebrates. Our rapidly advancing knowledge of invertebrate neurophysiology fosters the hope that a comprehensive account of the biology of intelligence is almost within reach. □

Geoffrey Hall is Senior Lecturer in the Department of Psychology, University of York, York YO1 5DD, UK.

Birth of the maser and laser

Shortly before his death in January 1984, Professor Alfred Kastler, winner of the Nobel Prize for physics in 1966 for his work on optical pumping, submitted to Nature a book review of Masers and Lasers by Mario Bertolotti¹. Published here is an edited and abridged version of the review.

THE word 'maser' was coined by Townes in 1951, as the acronym of 'microwave amplification by stimulated emission of radiation', and 'laser' was formed in the same way, 'light' replacing 'microwave'. In their famous *Physical Review* paper of 1958², Schawlow and Townes had used the term 'optical maser', a linguistic nonsense since it would literally mean 'optical microwave amplification...'. Incidentally, the most valuable property of these devices is not amplification, though they are widely used for this purpose in the microwave and optical regions, but self-sustained oscillation, so we should arguably call them 'mosers' and 'losers'!

The subject is, however, much older than the maser and the laser. Its prehistory began in 1916-17 with two papers by Einstein^{3,4} in which the distinction between spontaneous and induced emission of radiation was made. Both processes were shown to occur not as spherical waves, familiar from Maxwell's theory, but as a highly directional form of radiation (*Nadelstrahlung* rather than *Kugelstrahlung*, as Sommerfeld put it in 1922), the spontaneous emission having the same direction as the stimulating light and accentuating its intensity (negative absorption). The emission cannot, however, be in phase with the stimulus, for this does not permit conservation of energy; instead, the atomic emitters must be a quarter-period ahead of the driving field, as experimentally verified by Meslin. Einstein made no mention of the coherence of the stimulated emission, and it is tempting to speculate that this delayed the invention of the laser.

The exchange of momentum between atoms and light is clearly crucial and Einstein had examined this in great detail as early as 1909. As a working hypothesis, he suggested that light may be regarded as particles (our photons), which are to be thought of as local singularities in which the electromagnetic field energy is concentrated. These 'particles' are guided by the electromagnetic wave, interacting as they proceed. In Einstein's mind, therefore, wave properties such as interference and diffraction are a collective property of these particles. The richness of this idea in studies of both light and charged particles cannot be overstressed.

The important connection between the absorption and emission of light and the phenomenon of anomalous dispersion is perhaps not quite so obvious. The dispersion formula of Kramers and Heisenberg (1925)⁵ in fact contains 'negative dispersion' terms, which are connected with 'negative absorption' or, in other words, stimulated emission. The existence of these terms was demonstrated experimentally by Ladenburg and co-workers; in particular, Ladenburg and Kopfermann^{6,7} found that on raising the current density in a neon discharge, the negative dispersion terms grew more important, the temperature corresponding to the population ratio of the energy levels involved reaching values far higher than the kinetic temperature of the atoms and electrons. Had these authors continued their investigations and extended them to argon, they would probably have attained population inversion (a prerequisite for laser and maser action, to which I will return later). As Bertolotti mentions, Schawlow later thought it likely that Ladenburg and Kopfermann "did not continue their studies on anomalous dispersion because they believed so firmly in equilibrium that they thought it was impossible to go so far away from it as to have negative absorption"⁸.

In the 20 years that separated the 1930s—by which time the concepts of stimulated emission and negative absorption had become well established—and the invention of the maser and laser, the problem of magnetic resonance was solved and the process of optical pumping introduced. Although not directly relevant, the knowledge acquired was later useful in the development of lasers. Incidentally, Bertolotti's short note on the history of the classical theory of magnetism fails to mention Pierre Curie's thesis on the magnetic properties of matter, completed in 1895; he also wrongly credits Pauli with developing the concept of 'spin' in 1920 (see box), but the complex history of the different types of magnetism demands a fuller account than is given in the book.

In the early 1950s, prehistory ended and history began. The maser principle was developed independently at Columbia University (New York) by Townes, at the Lebedev Institute (Moscow) by Basov and Prokhorov and in the University of Maryland by Joseph Weber. It was at a conference in Ottawa in 1952 that Weber first mentioned publicly the principle of the maser, the text of his lecture appearing in print a year later⁹. The work of Townes and colleagues on the ammonia maser appeared in 1955¹⁰, a year after the first publication on the subject by Basov and Prokhorov, who, as Bertolotti notes, "pointed out the theoretical possibility of a device producing microwaves by using stimulated emission at an All-Union conference on radiospectroscopy in May 1952, but ... their first written paper was not published until October 1954"¹¹.

To generate coherent electromagnetic radiation, three conditions must be satisfied: stimulated emission, strong enough to surpass the absorption of the medium, must be obtained; a suitably inverted population must be created; and a convenient cavity must be used. In fact, the first two conditions are identical. As far as the last one is concerned, the cavity in the microwave region is just a suitably shaped metal box. The latter usually has the form of a cylinder closed at one end by a piston so that the proper mode of oscillation can be matched to the frequency of the atomic or molecular transition. Single-mode operation is easily achieved.

In the case of lasers, the appropriate cavity is a Fabry-Perot interferometer, as Schawlow and Townes pointed out in their fundamental paper of 1958², and if any French contribution to the laser deserves mention this is it. Fabry and Perot developed their instrument at the University of Marseilles in the early years of the century¹². The cavity of this device is cylindrical and the diameter of the two parallel circular faces is conventionally greater than their separation. Here I should digress briefly to say a word about R. Boulouch, who was a precursor of Fabry and Perot. In 1893 he published a description in the *Journal de Physique*¹³ of the very fine fringes he had obtained with thin silver layers, and mentioned that, for a suitable value of the distance between the reflecting surfaces, he had been able to separate the components of yellow sodium light: "Les anneaux dédoublés apparaissent, pour une épaisseur convenable de la lame, singulièrement nets" (for a suitable choice of the thickness of the plate, the two ring systems are remarkably distinct); he also noted that the line-shape could be explained in terms of the Airy function. Boulouch was at that time a grammar-school teacher at the Lycée of Bordeaux, where I myself taught in the years 1929-31, and where I met the elderly Boulouch, by then

Kastler's views on Pauli's contribution

BERTOLLOTTI is mistaken when he attributes the development of the concept of 'spin' to Pauli in his 1920 paper. This paper deals with the quite different problem of paramagnetism. 'Electron spin' was in fact discovered in 1925 by Uhlenbeck and Goudsmit. Why did this concept become necessary to atomic physicists? The extensive study of fine structure of spectral lines and the related anomalous Zeeman effect obliged physicists to introduce a new quantum number to stand alongside the principal quantum number n , which determines the energy of an atomic state, and the azimuthal quantum number k , which gives the angular momentum of a state, both in units of Planck's constant h and both being integers. (Today we use l where $l = k - 1$.) The third quantum number, which we represent by the letter j , sometimes takes a half integer value—this came as a surprise at the time—the vector representing the total angular momentum of the atom.

In a lecture given in Germany in 1965, when he was awarded the Planck medal, Samuel Goudsmit gave his personal version of the discovery of 'electron spin'. He cited a passage from a letter sent to him in March 1926 from Copenhagen by his friend and colleague L. H. Thomas. In this letter Thomas writes: "I believe that you and Uhlenbeck had the good luck to see your spinning electron published and discussed before Pauli knew of it. More than a year ago Kronig had conceived the rotating electron and had developed his idea; Pauli was the first person to whom he showed this paper. Pauli told him that it was a ridiculous assumption, and so it happened that the first person to see Kronig's paper was also the last one." And Thomas concludes the story by stating: "All this shows that God's infallibility is not extended to whom had nominated himself to be his vicar on earth." German

physicists have since composed the rhyme: "Der Kronig hatt 'den Spin entdeckt, Hatt' Pauli ihn nicht abgeschreckt." (Kronig would have discovered spin if Pauli had not discouraged him.)

There is another reason why Goudsmit adopted a very critical attitude towards Pauli. In his Nobel lecture of 1945 Pauli gave an explanation of the hyperfine structure of spectral lines by postulating the existence of a nuclear spin whose magnetic moment interacts with the moment of the electron cloud, an idea which he had published in 1924. He claimed that this publication had influenced Goudsmit and Uhlenbeck in their claim of an electron spin. Goudsmit maintained that Pauli had never considered an internal moment of the nucleus. He had especially not foreseen it for the proton. For him the nuclear angular momentum, with its small associated magnetic moment, was due uniquely to the orbiting electrically charged particles inside the nucleus, namely the protons and, as was assumed at that time, the electrons. So the discovery of electron spin as an internal momentum owed nothing to Pauli.

Pauli, in his young years, had been a severe judge for the physicists of his generation who feared his sharp criticism, and as we have seen his influence on the development of physics was not always a positive one. I made his personal acquaintance some years before his untimely death, meeting him at the Physics Summer School of Les Houches and in Varenna. In those days he had become a friendly and indulgent person. His exclusion principle, on which are built Fermi-Dirac statistics and the explanation of the shell structure of the electrons in an atom, remains a great achievement in the development of physics. □

retired. During 1889 and 1890, Charles Fabry had also taught at this same lycée, and had been a colleague of Boulouch, but he never mentioned the priority of Boulouch's researches. So far as I am aware, only Tolansky cites Boulouch's work, and he gives the wrong reference (1906 instead of 1893)¹⁴.

Let us now consider how population inversion was obtained. In Townes's first ammonia maser, molecules were excited into a higher energy state by means of a molecular beam, the molecules being separated by an inhomogeneous electrical field. Many celebrated experiments have used this atomic beam technique, but how many people remember that it was invented by Louis Dunoyer in Paris in 1919? Another elegant method of obtaining population inversion in the microwave region was exploited in Bloembergen's three-level maser. Some six different procedures have been used to trigger off laser action; it has often been found surprising that a simple gas discharge can create the necessary non-equilibrium state, as it does in the argon laser. Among the other methods, we recall electron impact, chemical reaction and Javan's proposal of 1959, which led to the first gas laser able to operate continuously, using two gases, and which was subsequently exploited in the high-power CO₂ laser. Bertolotti discusses all of these techniques thoroughly in Chapters 4, 5 and 6 of his book.

Finally, optical pumping can be used to obtain population inversion and it was in this way that Maiman operated the first laser, relying on the chromium ions in a ruby crystal¹⁵. Perhaps I can give my own version of the story. Optical pumping techniques were developed in my Paris research group during the period 1950–70 by a dozen young and enthusiastic research students under the guidance of Jean Brossel. It was this collaborative work that was the origin of my Nobel prize. In the very first paper in the new journal *Applied Optics*¹⁶, Francis Bitter surveyed this development and Jean Brossel has also given us his view of the story¹⁷. Ground-state optical pumping had been preceded by the double-resonance method, in which optical resonance produced by light is combined with magnetic resonance generated by a radiofrequency coil. As Bertolotti mentions, the Italian physicists Fermi and Rasetti had applied an oscillating magnetic field to atoms at optical resonance as early as 1925

and observed an oscillating Hanlé effect (depolarization of the resonant light). In his *Note e Memorie di E. Fermi*¹⁸, quoted by Bertolotti, Rasetti describes how

Fermi pointed out that, since the mercury resonance line showed an anomalous Zeeman effect with a Landé factor of 3/2, the mercury atom should more likely precess with a frequency 3/2 times higher than the Larmor frequency. The choice between the two alternatives might be decided by investigating the behaviour of the polarization under magnetic fields, of the intensity of about one gauss and a frequency of a few megacycles per second in approximate resonance with the precession frequency of the atom.

It thus seems that in 1925 the idea of Brossel's double-resonance experiment had already occurred to Fermi, but the fields applied (1–2 G) were insufficient to produce any appreciable resonance.

Brossel and I arrived at this method independently after an exchange of letters about the failure of attempts to investigate an excited state of mercury by Bitter's method; Brossel was working in Bitter's laboratory at Massachusetts Institute of Technology (MIT) at the time (1949). We published the principle together¹⁹, after which Brossel went on to perform the corresponding experiment and published the result. In my 1950 paper²⁰, I suggested altering the population of the magnetic m -states in the ground-state of the atoms by irradiating them with polarized light; if the polarization was linear, a change would be produced between states of large and small m (alignment), while if the polarization was circular (left- or right-handed), concentration of the atoms in either the negative or positive m -state (polarization magnetization of the ensemble of atoms) would be favoured.

As a young student, I had the opportunity of visiting Germany in 1922, where I bought Sommerfeld's book. I read it eagerly, being especially interested in the section on the "Auswahlprinzip and Polarizationsregel" ("selection principle and polarization rule") in which the thesis of Sommerfeld's Polish research student A. Rubinowicz is summarized. Optical pumping is just an application of this principle and rule.

In 1951, Brossel returned to Paris after six years abroad, three with Tolansky in Manchester and three with Bitter at MIT.

(British and American scientists may find it difficult to imagine the state of scientific isolation in France during the war years; we were cut off from the Allies and had no desire for contact with German scientists.) Brossel had a good idea of the progress of science abroad and knew what equipment we should need to resume our research, essentially oscillators and photomultipliers to develop the double resonance and optical pumping techniques. Our expenses were defrayed by the French Centre National de la Recherche Scientifique.

I am grateful to Bertolotti for recalling that we never intended to work on stimulated emission. Our aim was to change not the energy but the angular momentum of atoms, either in their ground state or in their excited states, thereby detecting magnetic resonance transitions which could be measured with high precision. My Nobel prize was given for the development of the optical methods of radiofrequency resonance (the 1964 Nobel prize in physics having been awarded to Townes, Basov and Prokhorov, the pioneers of the maser and laser). Despite this, my photograph appeared on the cover of a weekly journal with the caption "the father of the laser". I refuted this immediately in an address before the French Academy of Sciences, but there seems no way of suppressing this *canard* which continues to reappear intermittently. I deeply regret that Jean Brossel did not share the Nobel prize with me.

Bertolotti continues with an account of recent developments, the construction of the free-electron laser for example, and in

a concluding chapter examines the statistical properties of light, which are very different for coherent and incoherent radiation; this material is more demanding mathematically than the earlier parts of the book.

It is a pity that there was not room to give a more popular account of the applications of lasers and masers, which have become marvellous tools, in physics of course, but also in metrology, astronomy, medicine and industry, not forgetting the laser light that reads compact disks. The author alludes only briefly to such applications, observing that the field is "still open today"—perhaps the best applications are still to come? The role of the laser in turning Gabor's proposal of holography into a reality is briefly mentioned. Let us hope that the newer high-power lasers will never be used as weapons in outer space. Lastly, a word to the publishers: it would have made things much easier for the reader if the notes and references had been grouped at the end of the book, instead of having to be hunted for at the end of each chapter.

Let me not end on a negative note, however. Professor Bertolotti's book is a fascinating account of the history and prehistory of the maser and the laser. Most readers will enjoy his thumbnail sketches of the scientists involved, for his book tells us not only about physics but also about physicists, a very attractive feature. Although it is not aimed at the non-scientist, scientists from other fields will be able to follow the story since complicated mathematical reasoning is usually avoided.

1. Bertolotti, M. *Masers and Lasers: An Historical Approach* (Hilger, Bristol, 1983).

2. Schawlow, A. L. & Townes, C. H. *Phys. Rev.* **112**, 1940 (1958).

3. Einstein, A. *Mitt. phys. Ges., Zürich* **16**, No. 18, 47 (1916).

4. Einstein, A. *Z. Phys.* **18**, 121 (1917).

5. Kramers, H. A. & Heisenberg, W. *Z. Phys.* **31**, 681 (1925).

6. Ladenburg, R. & Kopfermann, H. *Z. Phys.* **48**, 26 and 57 (1928).

7. Ladenburg, R. & Kopfermann, H. *Z. Phys.* **65**, 167 (1930).

8. Schawlow, A. L. *2nd int. Conf. Laser Spectroscopy*, Haute Savoie, 23–27 June (1975).

9. Weber, J. *Trans. IRE prof. Grp. Electronic Devices* PCED-3, 1 (1953).

10. Gordon, J. P., Zeiger, H. J. & Townes, C. H. *Phys. Rev.* **99**, 1264 (1955).

11. Basov, N. G. and Prokhorov, A. M. *Zh. éksp. teor. Fiz.* **27**, 431 (1954).

12. Fabry, C. & Perot, A. *Annls Chim. Phys.* **12**, 459 (1897).

13. Boulouch, R. *J. Phys.* **2**, 316 (1893).

14. Tolansky, S. *Multiple Beam Interferometry of Surfaces and Films* (Clarendon, Oxford, 1948).

15. Maiman, T. H. *Nature* **187**, 493 (1960).

16. Bitter, F. *Appl. Opt.* **1**, 1 (1962).

17. Brossel, J. in *Polarisation, Matière et Rayonnement*, 143 (Prévies Universitaires de France, Paris, 1969).

18. Rasetti, F. *Note e Memorie di E. Fermi* Vol. 1, 159 (Accademia dei Lincei, Rome, 1962).

19. Brossel, J. & Kastler, A. *C.r. hebdom. Séanc. Acad. Sci., Paris* **229**, 1213 (1949).

20. Kastler, A. *J. Phys. Radium, Paris* **11**, 255 (1950).

High-power solid-state lasers

J. F. Holzrichter

Lawrence Livermore National Laboratory, Livermore, California 94550, USA

Laser power has increased by 11 orders of magnitude since Maiman's first demonstration of 10^3 W in 1960. Present high-power solid-state laser systems have been irradiating targets with 1–10 terawatts of power in 1-ns pulses at wavelengths varying from 1.05 to 0.26 μm . The Nova laser is beginning experiments in the 10 to 100-terawatt range. There are numerous possible means of extending solid-state technology to the 500 to 1,000-terawatt level and to efficient, high-average-power operation.

THE prospect of constructing lasers having peak powers greater than 10 TW excited the interest of experts in thermonuclear reactions in the early 1960s¹. This power, when focused to intensities of $\sim 10^{14} \text{ W cm}^{-2}$ onto small capsules containing deuterium and tritium fusion fuel, was thought to be capable of generating thermonuclear reactions^{2,3}. Other applications considered at the time included the optical pumping of X-ray lasers, and equation-of-state and other atomic-physics measurements at high temperature and density. Efforts to construct lasers powerful enough to demonstrate these concepts have been extremely successful (see Fig. 1): thermonuclear conditions in deuterium-tritium (DT) fuel are being attained⁴, stimulated emission at soft-X-ray wavelengths has been observed^{5,6}, and a variety of other measurements are being accomplished⁷.

Laser power has increased by 11 orders of magnitude, from Maiman's 1960 Q-switched ruby laser⁸, which produced ~ 1 kW, to the Nova laser, which can produce up to 100 TW (Fig. 2). In the 1960s the invention of mode-locking led to the generation

of laser pulses as short as 10^{-11} s, and Nd:glass gain media capable of amplifying these short pulses to moderate energies (~ 100 J) were developed. The appropriate combination of these technologies yielded lasers producing nearly 100 GW. The physics of these systems was incompletely understood, however, and while laser-target experiments were enticing, the results were erratic. By 1971, programmes were conceived to build 10-TW lasers to demonstrate inertial confinement fusion (ICF) concepts at spatial scales large enough to ascertain the potential for generating efficient thermonuclear burn⁹.

Inertial confinement fusion research

In 1971, three laser media—atomic iodine, CO_2 and Nd:glass—were available for scaling to laser sizes capable of producing 10 TW (that is, 10 kJ of energy in pulses shorter than 10^{-9} s). The atomic iodine laser system was not chosen for development in the national laboratory programmes in the United States because the diagnostics were somewhat difficult at the laser

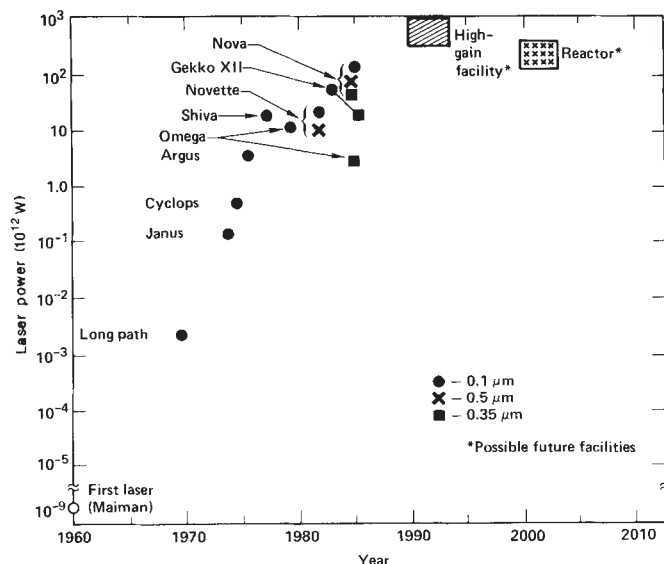


Fig. 1 Laser systems have increased dramatically in peak power since 1960. The laser names before 1980 are Livermore systems although other laboratories contributed significantly to this progress. Lasers in the 10-TW category were Shiva and Novette, the two-beam precursor to the 10-beam Nova system. Gekko XII, a 12-beam laser delivering 50 TW at $1\ \mu\text{m}$ or 20–50 TW at $0.5\ \mu\text{m}$, is beginning experiments at Osaka University. Omega is a 24-beam (12-TW, $1.0\ \mu\text{m}$) laser system at the University of Rochester, recently converted to 4 TW at $0.35\ \mu\text{m}$. The high-gain and reactor facility boxes show the objectives of high-power laser design.

wavelength of $1.3\ \mu\text{m}$ and because there was little experience in generating sub-nanosecond laser pulses⁷. The potential of the CO_2 laser to generate laser pulse energy with an efficiency of $\sim 10\%$, and its high-average-power capability for possible reactor applications, made it an attractive candidate. But the coupling of $10.6\ \mu\text{m}$ CO_2 laser light to dense target plasmas occurs at electron densities of $10^{19}\ \text{cm}^{-3}$, where electron-ion damping of plasma instabilities is weak, making the production of a very nonthermal plasma, unsatisfactory for inertial fusion applications, a distinct possibility. The attractiveness of the CO_2 laser system, and uncertainty over the seriousness of these plasma problems, led the group at the Los Alamos National Laboratory to develop CO_2 lasers having outputs of 10 TW (the Helios laser) and 30 TW (the Antares laser system) (ref. 9; see also similar

reports from the Los Alamos National Laboratory, 1971 to present).

At the Lawrence Livermore National Laboratory (LLNL), the Nd:glass laser was chosen as the major tool for research into the laser-plasma interaction¹⁰. Attractive features of the neodymium system are the ease with which pulse duration and pulse shape can be varied, the high optical-damage thresholds (which lead to moderately compact systems), the conveniently available beam diagnostics using film, silicon detectors or vidicons, and the scalability (through increased beam area or multiple beams) to high powers, which now exceed 100 TW and which can reach 1,000 TW. A variety of very efficient harmonic conversion techniques were also invented for this system, and these have been used to produce laser energy for a wide variety of experiments at 0.53 , 0.35 and $0.26\ \mu\text{m}$ (refs 11–13).

During the 1970s, experimenters used these systems and showed that both $10\text{-}\mu\text{m}$ and $1\text{-}\mu\text{m}$ laser wavelengths are too long for effective plasma heating. Studies using light at 0.5 , 0.35 and $0.26\ \mu\text{m}$ from harmonically converted Nd:glass lasers showed satisfactory plasma heating and confirmed earlier simulations. As a result, most ICF research is now being conducted with harmonically converted Nd:glass lasers. This research is being done in the United States at KMS Fusion Inc., Lawrence Livermore National Laboratory, the University of Rochester and the Naval Research Laboratory; in Europe at the Rutherford Laboratory and the Atomic Weapons Research Establishment, Aldermaston (both in England); at CEA Limeil in France; at the Lebedev and Kurchatov Institutes in Moscow; at the Max Planck Institute in Garching, FRG; and in Japan at Osaka University.

Research to provide other sources of short-wavelength light is also progressing. This work centres on the krypton fluoride laser¹⁴, with output at $0.249\ \mu\text{m}$, and the harmonically converted atomic-iodine laser^{7,15}, with outputs at 1.3 , 0.65 , 0.44 and $0.32\ \mu\text{m}$. These systems are being developed at several laboratories, but they have not yet been used for target experiments at multiterawatt levels.

While it is now clear that Nd:glass laser technology will suffice for the necessary laser-plasma research, it was thought necessary to invent a separate high-average-power, efficient laser technology for potential commercial applications. This process has been ongoing since the mid-1970s; there is not space here to review progress in this area. One surprising result of these studies, described in ref. 16, is the extent to which advanced solid-state systems, building on Nd:glass technology, may play a role.

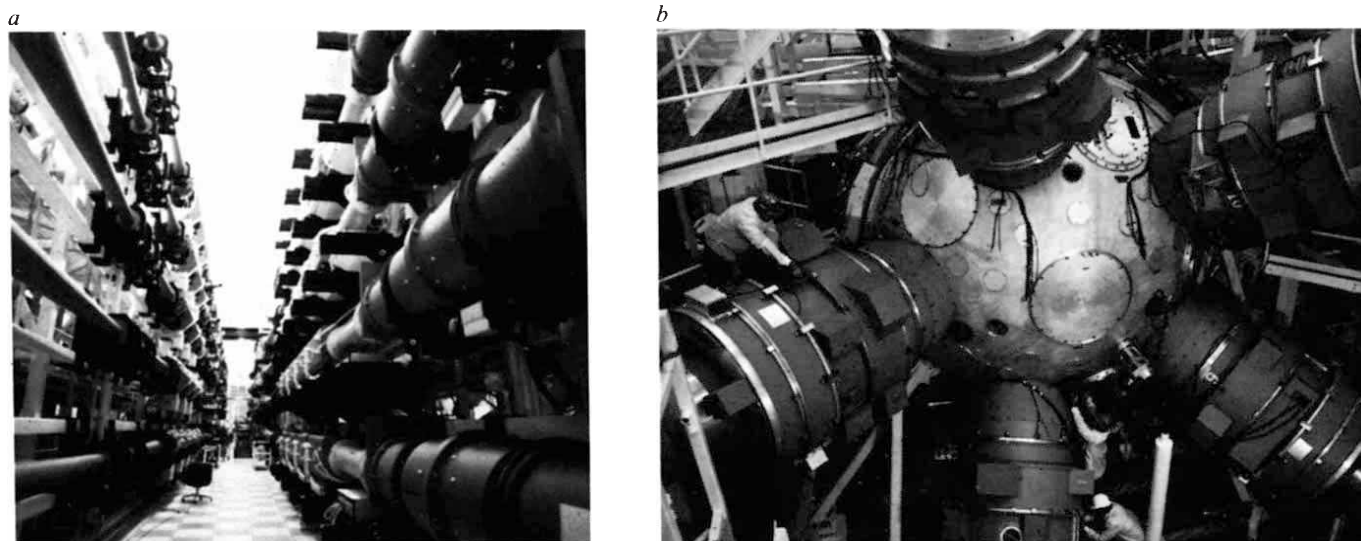


Fig. 2 *a*, The 10-beam Nova laser will provide 80–100 TW of peak power at $1.05\ \mu\text{m}$ and 40–70 TW of power at 0.53 or $0.35\ \mu\text{m}$ for fusion experiments. The pulse duration (and pulse shape) is variable from 70 ps to >10 ns. *b*, The target chamber. The large barrels attached to the target chamber contain the harmonic conversion arrays and the focusing lenses.

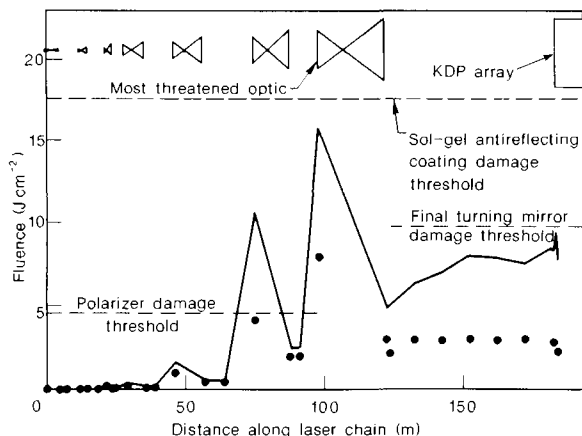


Fig. 3 The Nova laser beam is expanded, spatially filtered and re-imaged (projected) from a smaller amplifier section to the next larger section. Interchain peak (solid line) and average fluences (dots) as a function of distance along the laser chain, are shown for a 10.5-kJ laser pulse 1 ns long. The sizes and locations of the spatial filter lenses are indicated along the top of the figure. Entrance-lens surfaces will suffer damage if the peak fluence exceeds the indicated limits.

Laser physics and design

The physical constraints that control laser operation at 10 TW are reasonably well understood, and recently constructed systems have been designed to accommodate them¹⁷. These constraints are associated with the generation and shaping of the initial laser light pulse, the production and control of gain (and loss), the control of linear and nonlinear propagation along the beam path to keep the beam highly focusable, and the harmonic conversion of the high-power fundamental laser wavelength. Additional engineering concerns include cost, maintenance, and the diagnostics and control of the laser-target system.

The design of present Nd:glass lasers aims at realizing gain and propagation performance as close as possible to physical limits. While we use the Nova fusion laser as an example¹⁸, the design of high-power laser systems based on the other media is subject to similar constraints.

The amplification of a 10^{-4} -J, 1-mm-diameter oscillator pulse to a $\sim 10^4$ -J, 74-cm beam is accomplished by paying careful attention to the optical properties of the laser media and the optical system. Astronomical telescopes are used to enlarge and project the laser beam from the oscillator to the target-focus lens. The telescopes also project and image the beam shape from one stage to the next down the chain¹⁹; this prevents the beam shape from diffracting and developing highly modulated

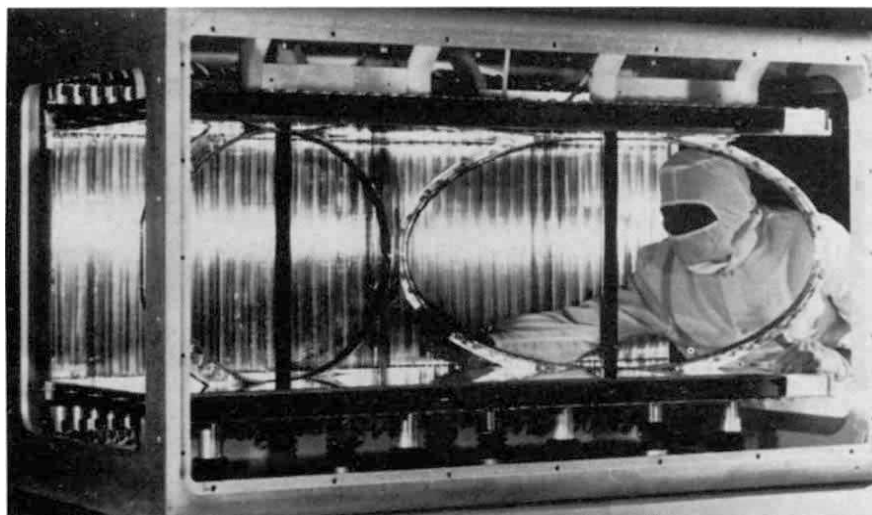
diffraction 'ripples' as it propagates. That is, a well-shaped (apodized) beam that effectively fills the laser aperture can be magnified, projected through each laser amplifier where it is intensified, then further magnified and amplified in the next stage until it reaches the desired energy (consistent with the optical-damage threshold of the final element and with nonlinear self-focusing limits), and projected and amplified again (see Fig. 3). The telescope also serves a second important function, that of spatial filtering¹⁹. At the focus of the telescope system the laser beam is passed (in vacuum) through a pinhole that removes the high-frequency spatial noise from the beam. A second telescope lens then recollimates the expanding beam and projects a 'smoothed' image of the input aperture a distance Mf further downstream, where M is the magnification and f the length of the spatial filter. (A consequence of this approach is that the long distance travelled by the beam in Nova-like lasers is mostly through telescopes.)

The laser gain is determined by the stimulated-emission cross-section of the laser medium, by the pumping-rate limits set by the power supply, and by one of two gain-limiting processes. The first of these processes is stimulated depumping by amplified spontaneous emission (ASE) and the second is feedback oscillation (usually unexpected) from reflecting sources such as metal lens holders, screwheads, or the walls of the amplifier case. Stimulated depumping occurs when a spontaneously emitted photon traverses the gain medium and stimulates other ions to de-excite, thus depumping the medium. When the 'gain-coefficient $\times$ gain-length product (αl)' in the laser medium exceeds 4.5, this rate approaches 30% of the pumping rate for usual pumping situations. A greater loss rate than this is undesirable. Furthermore, at such a gain level, $G = \exp(4.5) = 90$, a 1.1% or larger reflection can cause a parasitic oscillation to start. This is avoided in glass lasers by cladding the edges with an index-matching absorbing glass yielding $<10^{-3}$ reflectivity. In the largest Nova amplifier (see Fig. 4), the laser plates are also cut and clad along the middle to control the transverse gain path and thus the ASE loss.

The efficiency of gain production is a complex function of the power-supply to gain-medium transfer efficiency (typically 1%) and the gain-medium to laser-light transfer efficiency (typically 25%). In research lasers, system efficiency is important only insofar as it (and the per-unit cost of power-supply) contributes to the total system cost. The very low cost of flashlamps and discharge capacitors compensates for the present low efficiency to provide acceptable power supply costs. However, for commercial applications such as the production of electricity or nuclear fuel, system efficiencies near 10% would be necessary²⁰.

Gain is developed when Nd^{3+} ions in optical glass plates are pumped by xenon flashlamps for 600 μs duration in an amplifier structure (Fig. 4). The gain coefficient α in the glass in the larger

Fig. 4 The neodymium glass plates are mounted at Brewster's angle with respect to the laser beam to provide zero reflection loss. Light from the flashlamps excites the Nd^{3+} ions to produce gain.



a

b

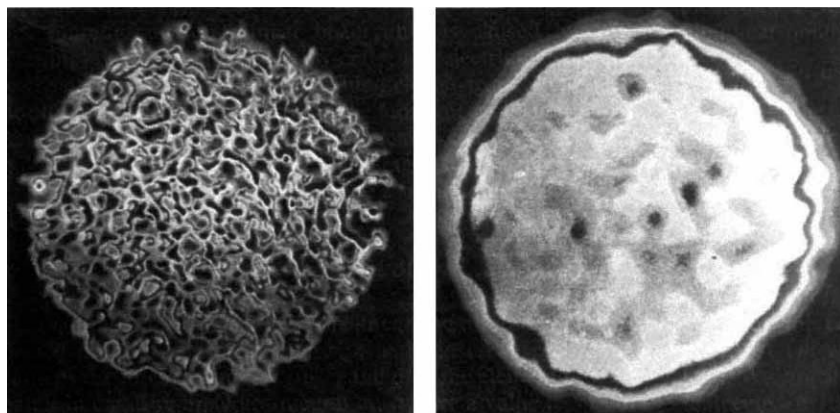


Fig. 5 High-power laser beams will self-focus when propagated at high powers through long glass paths. The criterion is given by the *B*-integral and by the spatial noise sources in the laser system. When spatially filtered, the beams are smoothed so that further propagation is possible. *a*, Into the spatial filter; *b*, exit from the spatial filter.

amplifier is typically 0.05 cm^{-1} , so that the beam must traverse a total path of 45 cm in the gain medium to reach a gain of 10. Ten 4-cm plates set at Brewster's angle in a series of disk amplifiers like those shown in Fig. 4 are used to provide the gain path in this amplifier stage. Six stages of amplification are used in each beam line to amplify the beam from 1 J at the entrance to 10^4 J at the exit.

As the beam propagates down the laser chain, its phase front is modified by linear and nonlinear phase aberrations. Linear aberrations cause small modifications to the phase front that reduce the focusability of the beam on the target. The optical glass in which the neodymium ions are dissolved can be made homogeneous to 2 parts in 10^6 , which would yield approximately 2 waves of aberration per metre of path in glass. Present polishing techniques are capable of compensating for these aberrations: when the beam passes through the 150 surfaces and 3 m of glass in a Nova laser beam, an acceptable 2–3 waves of total linear aberration result.

The most nonlinear (and permanent) impediment to beam propagation is optical damage. The energy fluence in the beam, $5\text{--}10 \text{ J cm}^{-2}$, is enough to cause the explosion of small absorbing inclusions in coatings or within laser glass or lenses, obscuring part of the beam and permanently damaging the laser component. Laser glass and other transmitting optics can be made inclusion-free², which will allow them to sustain intensities greater than $10^{10} \text{ W cm}^{-2}$ in 1-ns pulses. However, to obtain good production yields of inclusion-free glass for reliable operation at these fluence levels, additional work is required. A great deal of effort has been expended on developing coatings that can withstand operation at high fluence. Particularly successful anti-reflection coatings based on 'sol-gel' coating technology and on 'neutral-solution' surface etching processes have been produced and used in the Nova laser²¹. These coatings are virtually defect-free and exhibit damage thresholds three times higher than those produced using electron-beam deposition procedures.

A more gradual nonlinear beam limit is caused by self-focusing¹⁷. At intensities $>10^9 \text{ W cm}^{-2}$, the electric field of the laser beam exceeds 10^6 V cm^{-1} . This field is strong enough to modify the electron orbits around the atoms or molecules in the medium through which the beam propagates, changing the index of refraction. This causes a self-focusing instability to grow at a rate proportional to $\exp(2B)$, where *B* is the so-called breakup or *B*-integral

$$B = 0.026 \frac{n_2}{n\lambda} \int I \, dl \quad (1)$$

Here *n* and *n*₂ are, respectively, the linear and nonlinear indices of refraction in electrostatic units, λ is the laser wavelength in μm , *I* is the intensity in GW cm^{-2} , and *dl* is the element of path length in cm. The breakup integral must have a value less than ~ 3 for propagation to occur without severe loss. The nonlinear index of refraction changes the index of refraction by $\sim$ one

part in 10^6 at peak intensity. This change is small (it corresponds to less than one wave of beam phase retardation per metre of glass path traversed). But if the transverse dimension over which it takes place is small ($\sim 1 \text{ mm}$), the induced focal power of the resulting 1-mm-diameter 'lenslet' will cause a 1-mm portion of the beam to self-focus, and the corresponding energy will not be available for harmonic conversion or for focusing on a target (see Fig. 5). Nonlinear beam self-focusing was the chief cause of nonreproducibility in early laser-target experiments. In extreme cases, the entire laser beam self-focused into thousands of small light filaments that could not be focused on the target and which 'drilled' small holes in the laser glass, the mirror coatings and the lens surfaces. Control of nonlinear self-focusing made Nd:glass lasers useful for target irradiation and made them scalable to much higher power than was at first predicted.

To provide adequate gain and to control self-focusing, neodymium-doped phosphate glass was selected as the gain medium. Its gain characteristics are compatible with 5-J cm^{-2} output fluence, and its self-focusing characteristics permit the propagation of twice as much optical fluence as in silicate laser glasses. This permitted propagation at higher fluences through longer gain-medium paths than in previous fusion lasers, such as the Shiva laser at LLNL. The longer gain path is needed to amplify the laser pulse to the 5-J cm^{-2} (5-GW cm^{-2}) level at 1-ns pulse duration, or to 10 J cm^{-2} when the system is used for experiments with pulse durations $>3 \text{ ns}$. The lower nonlinear index of refraction of the phosphate glass permits high-power operation with very short pulses ($\sim 0.1 \text{ ns}$) also.

Harmonic conversion

Harmonic conversion is used to generate shorter wavelengths for target irradiation. This is feasible because of the high phase uniformity of high-power beams from modern lasers, and because the power density in such beams is high enough (GW per cm^2) to drive a nonlinear polarization in the atoms or molecules of non-centrosymmetric crystals. The magnitude of the linear polarization is

$$P = \chi \epsilon_0 E = 10^{-4} \text{ debye per molecule per GW cm}^{-2} \quad (2)$$

where χ is the susceptibility, ϵ_0 is the vacuum permittivity, and *E* is the electric field of the light wave. If a non-centrosymmetric crystal is used, the restoring force on the electron cloud is not symmetric about the equilibrium position, and the susceptibility can become proportional to *E*:

$$\chi = \chi_0 - \chi_1 E + \dots \quad (3)$$

When equation (3) is substituted into equation (2), there results a term in the polarization, $\chi_1 \epsilon_0 E^2$, that leads to an oscillating dipole at twice the fundamental frequency ω_0 of the electric field. A typical $2\omega_0$ nonlinear polarization is 10^{-9} debye per molecule, or 10^{-5} times the linear polarization.

One can imagine this nonlinear polarization developing in the first plane of atoms in the crystal as the beam strikes the

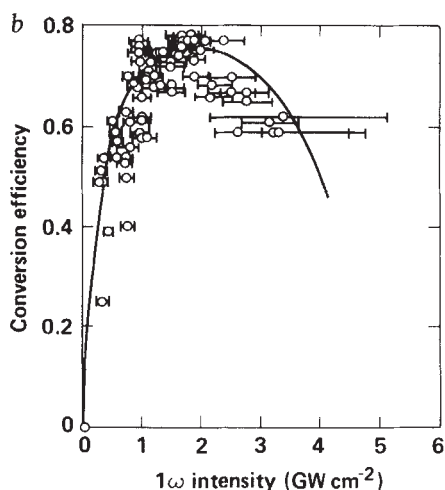
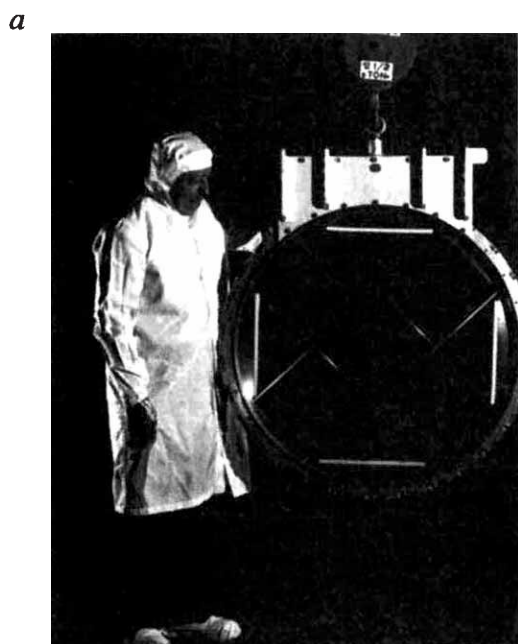


Fig. 6 High-power, 1.05- μm laser beams are sufficiently uniform in phase to uniformly convert to 0.53, 0.35 or 0.26 μm with high efficiency. *a*, A 3×3 array of 1.8-cm-thick KDP harmonic conversion crystals. *b*, Conversion efficiencies with a similar array on the Novette laser.

surface. We can arrange for the radiated waves from each plane of atoms to add coherently to those of the succeeding planes, so as to generate a strong harmonic wave; we do this by choosing the propagation direction in uniaxial or biaxial crystals so that the phase velocities of the harmonic and fundamental waves are the same (the waves are then said to be phase-matched). Phase-matching allows for very efficient harmonic conversion when the phase front of the incoming fundamental laser light is uniform enough to maintain these conditions over the entire beam area and throughout the length of the crystal. For example, with a properly oriented 1.8-cm-thick crystal of potassium dihydrogen phosphate (KDP), we can convert 1.05- μm laser light at 2 GW cm^{-2} to the second harmonic at 0.53 μm with up to 80% efficiency¹¹⁻¹³. When losses in the optical components (lenses, diagnostics, blast shields) of a fusion laser beam line are taken into account, it is possible to deliver over 50% of the laser's 1.05- μm energy to a target at 0.53 μm ; 70% conversion efficiency may be expected with additional engineering improvements^{22,23} (Fig. 6).

By using a second crystal after the first crystal, efficient conversion to the third harmonic at 0.35 μm or the fourth harmonic at 0.26 μm is possible. For the third-harmonic process, the

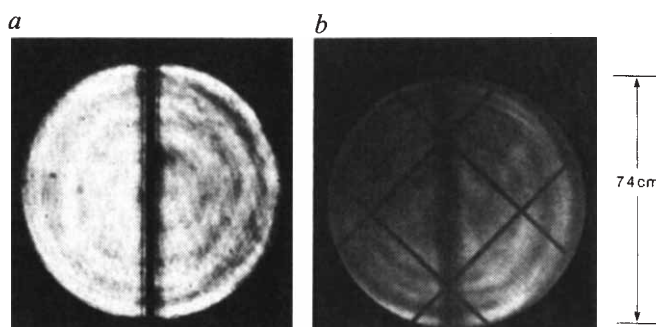


Fig. 7 Amplitude pictures of 74-cm-diameter beams on Novette. The quality of these beams permits focusing to 100- μm -diameter focal spots, leading to power densities $>10^{17} \text{ W cm}^{-2}$ on target. However, most target irradiation is conducted at 10^{14} – $10^{15} \text{ W cm}^{-2}$. The dark vertical line in the middle of the beam is due to the gain-parasitic absorber in the largest amplifier. The 3×3 checked pattern is due to the array of harmonic conversion crystals. *a*, 10 TW, 1.05 μm ; *b*, 10 TW, 0.53 μm .

second crystal is oriented to phase-match a one-to-one photon mix of two input waves, one at 0.53 and the other at 1.05 μm , which beat together through the χ_1 term in equation (3) to produce the output wave at 0.35 μm . Experiments at Rochester and at Lawrence Livermore have demonstrated third-harmonic conversion efficiencies of up to 80% (refs 11–13). For the fourth-harmonic process, the second crystal is oriented to convert a 0.53- μm wave from the first crystal to 0.26 μm .

The Nova laser

The Nova laser was designed to be a versatile experimental system capable of irradiating targets with pulse durations from 0.05 ns to over 10 ns at wavelengths of 1.05, 0.53 and 0.35 μm (ref. 24). Experiments at 0.25 μm can be accomplished with some modification. In 1982, early production of two of Nova's 10 beams were configured into a facility called Novette^{22,23}, which was used to verify the Nova laser design and the process of harmonic conversion to 0.53 μm , and to verify models of plasma heating by 0.53- and 0.25- μm laser light. This work was largely successful²⁵. The objective of the full 10-beam Nova laser is to extend these results and produce conditions of DT-fuel compression and laser-plasma coupling similar to those required by high-gain targets.

The diameter of Nova's beams posed special problems for harmonic conversion. Our solution was to assemble slices cut from 27-cm-wide KDP crystals into a 3×3 mosaic array with a 74-cm-diameter aperture (see Fig. 6). A significant technological innovation was the use of a diamond cutting tool on a precision lathe to cut each crystal surface to the required smoothness and flatness and to the correct angle with respect to the crystal axes. In addition, it was necessary to develop coatings, mechanical supports and diagnostic techniques. The final harmonic system is remarkably simple, using advanced sol-gel coatings and a two-crystal design that efficiently generates 0.53 μm or 0.35 μm light with a simple tilt ($1/4^\circ$) and rotation (10°) of the array. A singlet-lens focusing system, using chromatic dispersion in the lens, together with a central beam block, can focus the beam at the fundamental 1.05- μm wavelength, or at the two harmonic wavelengths, without other, unwanted wavelengths reaching the target.

At the highest beam powers delivered by Nova, stimulated nonlinear optical processes can occur: stimulated Raman scattering from nitrogen in the beam path, additional nonlinear self-focusing in the air path, and transverse stimulated scattering in Raman-active materials (for example, the KDP crystals) were the expected possibilities. Careful measurements of the thresholds for these processes gave us confidence that the system would perform satisfactorily at the desired powers. With the Nova beams finished, we have now measured a gain of 2.7 cm per TW for stimulated Raman scattering in a very long beam

path²⁶. This gain is high enough to limit Nova's output power, so we are planning to remove the nitrogen from the beam tube. Nova's output in 1-ns pulses will be 80–100 kJ at 1.05 μm , with better than 50% conversion to 0.53 or 0.35 μm , depending on the pulse duration. Figure 7a, b shows the 1.05- and 0.53- μm beam images, respectively, from a Nova beam when it was used on the Novette laser.

Future developments

The future development of high-power solid-state lasers will follow two directions. A 5–10-MJ, 1,000-TW laser will be required to demonstrate and optimize the performance of high-gain ICF targets (see Fig. 1). A high premium will be placed on the versatility and cost of this system technology, and on the utility of the technology in future industrial, defense and ICF applications. The second direction is toward high average power and very efficient operation at moderate cost. Systems with these characteristics would be required for electrical power generation or for production of nuclear fuel.

Work on both types of systems is progressing. Considering our views in the early 1970s, we are surprised that the scale of the Nd:glass laser, as well as those of the I, KrF and CO₂ lasers, can indeed be increased so as to reach the energy and power

levels necessary for high-gain fusion demonstrations. The harmonically converted glass laser appears to offer the most flexible and economical approach to meeting this objective. This conclusion is based on the physics constraints of the competing systems and on the engineering and manufacturing experience developed in building large CO₂ and glass fusion-laser systems and smaller experimental KrF laser systems. The prospect of high-repetition-rate (5–10 Hz), highly efficient (>10%) solid-state lasers using crystalline hosts rather than glass hosts and using solid-state pumping rather than flashlamp pumping has been theoretically supported by Emmett *et al.*¹⁰ and is in the process of experimental demonstration.

I thank my colleagues in the Laser Fusion Program at the Lawrence Livermore National Laboratory who have worked to design, construct and use these systems. In particular, J. L. Emmett has been responsible for the overall programme direction. W. F. Krupke developed many advanced laser concepts, J. B. Trenholme, W. F. Hagen and W. E. Warren developed much of our design methodology, J. T. Hunt and D. R. Speck are readying Nova for target experiments, R. O. Godwin and W. W. Simmons designed and managed Nova construction, K. R. Manes and G. J. Suski led the team that built and used Novette, the precursor to Nova, and D. Eimerl and M. A. Summers provided many of the harmonic conversion ideas.

1. Basov, N. G. & Krokhin, O. N. in *Proc. 3rd Quantum Electron. Int. Conf.*, Paris, Vol. 2, 1373–1377 (Columbia University Press, 1964).
2. Nuckolls, J. H., Wood, L. L., Thiessen, A. R. & Zimmerman, G. B. *Nature* **239**, 139–142 (1972).
3. Emmett, J. L., Nuckolls, J. H. & Wood, L. L. *Scient. Am.* **230**, 24–37 (June 1974).
4. Ze, F. *et al. Nucl. Fusion* (in the press).
5. Matthews, D. L. *et al. Phys. Rev. Lett.* **54**, 110–113 (1985).
6. Rosen, M. D. *et al. Phys. Rev. Lett.* **54**, 106–109 (1985).
7. Schwarz, H. J. & Hora, H. *Laser Interaction and Related Plasma Phenomena*, Vols 1–6 (Plenum, New York, 1971–84).
8. Maiman, T. H., Hoskins, R. H., D'Haenens, I. J., Asawa, C. K. & Evtuhov, V. *Phys. Rev.* **4**, 1151–1157 (1961).
9. *Laser Program a. Rep. Lawrence Livermore National Laboratory 1971–84* (UCRL-50021-71 to -84).
10. Emmett, J. L., Krupke, W. & Davis, J. I. *IEEE J. quant. Electron.* **20**, 591–602 (1984).
11. Seka, W. *et al. Optics Commun.* **34**, 469–473 (1980).
12. Craxton, R. S. *Optics Commun.* **34**, 474–478 (1980).
13. Linford, G. J. *et al. Appl. Opt.* **21**, 3633–3643 (1982); **22**, 1957–1958 (1983).
14. Rhodes, C. K. (ed.) *Excimer Lasers* 2nd ed (Springer, New York, 1984).
15. Brederlow, G., Witte, K. J., Fill, E., Hohla, K. & Volk, R. *IEEE J. Quantum Electron.* **QE-12**, 152–155 (1976).
16. Emmett, J. L., Krupke, W. F. & Trenholme, J. B. *Physics of Laser Fusion. The Future Development of High-Power Solid State Laser Systems* Vol. 4 (Lawrence Livermore National Laboratory UCRL-53344, November 1982).
17. Holzrichter, J. F. *et al. J. Fusion Energy* **2**, 5–45 (1982).
18. Simmons, W. W. *et al. in Proc. 9th Symp. Engng Problems of Fusion Res.*, Chicago, October 1981, Vol. 2, 1221–1273 (IEEE Publ. No. 81CH1715-2 NPS).
19. Hunt, J. T., Glaze, J. A., Simmons, W. W. & Renard, P. A. *Appl. Opt.* **17**, 2053–2057 (1978).
20. Nuckolls, J. H. *Phys. Today* **35**, 24–31 (1982).
21. *Proc. Symp. on Laser Induced Damage in Optical Materials* (National Bureau of Standards Spec. Publ. 1976–84).
22. Manes, K. *et al. Lawrence Livermore National Laboratory Rep. UCRL-90054-Rev. 1* (1984).
23. Manes, K. R. & Simmons, W. W. *J. opt. Soc. Am.* **A2**, 528–538 (1985).
24. *Lawrence Livermore National Laboratory Rep. UCRL-52000-85-2* (February 1985).
25. Holzrichter, J. F., Campbell, E. M., Lindl, J. D. & Storm, E. *Science* (in the press).
26. Hennesian, M. A., Swift, C. D. & Murray, J. R. *Lawrence Livermore National Laboratory Rep. UCRL-92621* (May 1985).

Laboratory production of X-ray lasers

M. H. Key

Laser Division, Science and Engineering Research Council Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK

Laser-produced plasmas seem to be the most promising sources of soft X rays at the high intensities required for laser action. Strong laser amplification has recently been demonstrated at the lowest X-ray energies, suggesting that true soft X-ray lasers will soon be available in the laboratory.

A MAJOR objective in laser research is to develop X-ray lasers¹. Optical lasers revolutionized optical methods because of the high power, collimation, spectral purity and coherence of their beams. Similar, dramatic consequences could be expected if X-ray lasers became widely available.

Many important scientific applications of intense X-ray emission are already being developed using synchrotron radiation sources². These emit incoherent X-rays from circulating bunches of highly energetic (GeV) electrons deflected by the magnetic field in an accelerator storage ring. A pure continuum of frequencies from infrared to X-ray is generated with high average power in subnanosecond pulses at high (100 MHz) frequency. Monochromators are used to select particular frequencies for applications such as diffraction or spectroscopy, and the emission pulses are accumulated over periods from seconds to hours to obtain sufficient signal.

More recently, single-pulse X-ray measurements have become possible using laser-produced plasmas³. Their incoherent emission results from thermal collisions of electrons with ions in a hot plasma produced by focusing a powerful laser pulse on to a solid target. A continuum of frequencies is emitted by free electrons and narrow lines are emitted by electrons bound to ions. The single-pulse spectral brightness is up to 10⁶ times greater than for synchrotron radiation sources as illustrated in Fig. 1.

To produce coherent laser X-rays has proved more difficult. For a constant laser amplification coefficient, the required fluorescent emission per unit volume on the laser transition scales as frequency $\nu^{4.5}$. Hence there is a need for exceptionally intense X-ray sources for X-ray lasers, the most promising of which seem to be laser-produced plasmas. In this context, a recent experiment at the USA Lawrence Livermore National

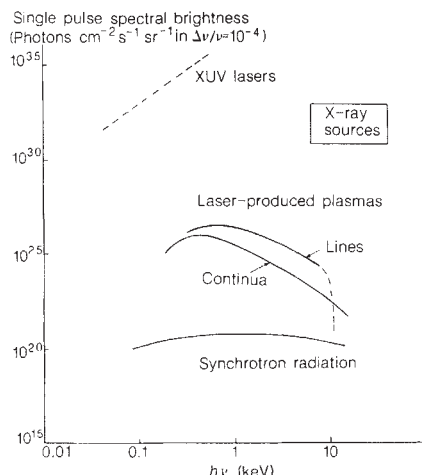


Fig. 1 Comparison of the absolute spectral brightness of XUV and X-ray sources. (Synchrotron data are from the SERC Synchrotron Radiation Source (SRS) and laser-produced plasma data are compiled for 10^{12} watt Nd glass lasers.)

Laboratory (LLNL) has generated much excitement by convincingly demonstrating strong laser amplification at a wavelength of 20.6 nm (photon energy $h\nu = 60$ eV; where h is Planck's constant) in a laser-produced plasma⁴. This is the most significant recent advance in short-wavelength lasers, the previous limit having been set at 116 nm ($h\nu = 10.7$ eV) in 1972 with the discovery of the Werner band H_2 laser. Although 116 nm has been the limit for laser action, generation of the seventh harmonic of ultraviolet laser light in gases has extended the short wavelength limit for high-power coherent radiation to 30 nm⁵.

It should be emphasized, however, that 60 eV is only an extreme (X) UV photon energy; a true soft X-ray laser will require a further order of magnitude increase in photon energy. However, the significance of the recent progress at LLNL is that it gives more confidence that a new family of XUV lasers based on high temperature laser-produced plasmas can be developed in the XUV/soft X-ray region >60 eV. Here I describe the characteristics of these lasers, explain how they operate, review recent progress and speculate on their future applications.

Background

All lasers have in common the fact that they produce light amplification by stimulated emission of radiation, hence their name. Stimulated emission was first postulated by Einstein as the quantum inverse of absorption. A quantum of electromagnetic wave energy (or photon), whose energy is $h\nu$, can resonantly stimulate the emission of an identical photon by an atomic electron in an excited state if the energy of the excited state E is $h\nu$, greater than that of a lower state. The new photon is coherent with the stimulating photon, that is the electromagnetic waves have the same phase. Stimulated emission amplifies electromagnetic waves and the inverse process of absorption attenuates them.

The only difference between an XUV laser and a more usual visible light laser is that the laser transition is an XUV frequency. Such transitions occur in high temperature plasmas of highly charged ions and electrons. The reason can be seen in Bohr's elementary model of a hydrogen-like ion illustrated in Fig. 2. The positive nucleus, with one orbiting electron, has a system of energy levels in which the transition energies are

$$h\nu = Z^2 h\nu_H \quad (1)$$

where Z is the charge number of the nucleus and $h\nu_H$ is the transition energy in a hydrogen atom ($Z = 1$).

For example, the Balmer α transition illustrated in Fig. 2 emits XUV photons ($h\nu \sim 68$ eV) when $Z = 6$. The hydrogen-like ion is produced by removing five electrons from a carbon atom leaving a positive nucleus with a single electron, which is isoelectronic (having the same number of electrons) as a hydrogen

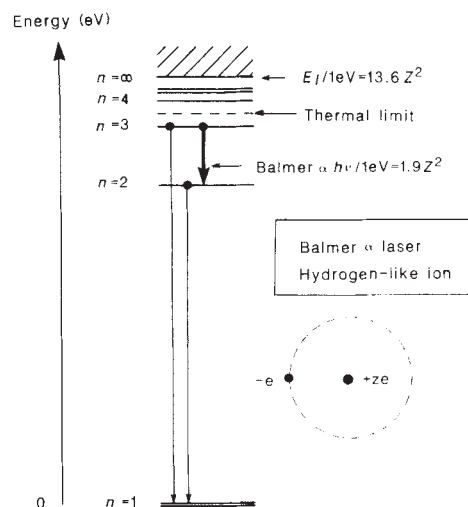


Fig. 2 Energy levels of a hydrogen-like ion of nuclear charge $+Ze$ showing the location of the thermal limit when population inversion is obtained between $n = 2$ and $n = 3$.

atom. Such isoelectronic scaling to higher Z shifts the emitted frequency towards the X-ray region.

Production of highly charged ions in plasmas is described by a well-known equation due to Saha and requires high temperatures such that the energy kT in collisions between electrons and ions is an approximately constant fraction f of the ionization energy $Z^2 E_H$

$$kT = fZ^2 E_H \quad (2)$$

The net amplification coefficient resulting from the balance between absorption and stimulated emission on any atomic transitions can be expressed as

$$G = (n_2 - n_1(g_2/g_1))(\lambda^2/8\pi)(A_{21}/\Delta\nu) \quad (3)$$

where n_1 and n_2 are number densities of the lower and upper states of degeneracy g_1 and g_2 , respectively, and A_{21} is the Einstein A coefficient or probability per unit time of a spontaneous transition from the upper to lower state of wavelength λ .

The spectral width $\Delta\nu$ of the transition due to the Doppler shift arising from the thermal velocity of the ions scales as

$$\Delta\nu/\nu \propto (kT)^{1/2} \text{ implying } \Delta\nu \propto \nu^{3/2} \quad (4)$$

There is also a fundamental frequency scaling of the Einstein A coefficient in an isoelectronic sequence

$$A_{21} \propto \nu^2 \propto \lambda^{-2} \quad (5)$$

Amplification (that is, negative absorption) is obtained with population inversion (excess of excited states), that is,

$$n_1(g_2/g_1) < n_2 \quad (6)$$

Equations (3) and (5) show that to sustain a given amplification requires a minimum density of excited states

$$n_2 \propto \Delta\nu \quad (7)$$

The spontaneous fluorescent emission from these excited states has a power per unit volume,

$$p = n_2 A_{21} h\nu \quad (8)$$

and it follows from equations (4) to (8) that

$$p \propto \nu^3 \Delta\nu \propto \nu^{9/2} \quad (9)$$

and it is this $\nu^{9/2}$ scaling of the fluorescent power that makes it very difficult to achieve laser action at high frequencies.

Another major problem is to find methods of creating population inversion when nature tends to create thermal equilibrium in which the well-known Boltzmann relations apply, giving no amplification at any temperature, that is

$$n_2(g_1/g_2) = \exp(-h\nu/kT)n_1 \quad (10)$$

Thermal equilibrium is a consequence of thermodynamic detailed balance in which each transition inducing process is balanced by its inverse process, for example



here collisional ionizations are balanced with three-body recombination. Collisional processes are always balanced by their inverse and therefore in the limit of high plasma density, where collisions are dominant, thermal equilibrium level populations prevail. But spontaneous radiative transitions are most rapid across large energy gaps ($A_{21} \propto \nu^2 \propto Z^4$), whereas for collision induced transitions (of rate $K_{21}n_en_i$, where n_e and n_i are the number densities of electrons and ions) the opposite is true, with the rate coefficient K_{21} scaling⁶ as

$$K_{21} \propto \nu^{-1} kT^{-1/2} \propto \nu^{-3/2} \propto Z^{-3} \quad (12)$$

The result is that we can identify, in a plasma of lower density, a 'thermal limit' as illustrated in Fig. 2. Above the limit, where energy levels are closely spaced, collisions dominate and below the limit radiative rates are greater than collisional rates. Radiative processes are not in detailed balance unless spontaneously emitted photons are reabsorbed by the plasma. This will not occur unless the plasma volume is large. Thus levels below the thermal limit need not be in thermal equilibrium and, with suitable tricks, population inversion can be created, for example, between levels of principal quantum number $n = 3$ and $n = 2$ in Fig. 2.

In isoelectronic scaling of laser mechanisms to shorter wavelengths it is generally necessary to preserve the balance between radiative and collisional processes so that the thermal limit remains at the same principal quantum number. From the preceding discussion it follows that the temperature $kT \propto Z^2$ and the particle number density $N \propto Z^7$ requiring plasmas of progressively higher temperature and density for lasers of higher frequency.

XUV laser characteristics

The new family of XUV lasers will probably be amplified spontaneous emission (ASE) lasers using single or double transit amplification of spontaneous emission in a long thin volume of laser medium as illustrated in Fig. 3. They are unlikely to have the usual two mirror laser resonator because of the combined problems of short duration of amplification, mirror damage, difficulty of output coupling and criticality of alignment. Here I outline the performance characteristics of such lasers.

The single transit amplification is $\exp(GL)$, where G is the amplification coefficient and L the length. The ASE laser requires a huge single transit amplification to reach the saturation condition where stimulated emission becomes the dominant process. Any plasma-based laser would require $GL \geq 14$, that is, $\exp(GL) \geq 10^6$.

The angular divergence of the laser beam is simply the width to length ratio w/L of the gain medium for a single transit (or the potentially much smaller width to overall length ratio w/L_0 for a double transit with a single mirror), as shown in Fig. 3 and will typically be 10^{-3} – 10^{-2} rad. Further simple atomic physics and plasma physics arguments show that the saturated output power of the laser depends mainly on the frequency and is proportional to the beam area. Assuming a beam area of 10^{-4} cm², typical of the cross-sectional area of the plasma (Fig. 3c) the output power is in the megawatt range

$$P = (h\nu/100 \text{ eV})^{4.5} \text{ MW} \quad (13)$$

The duration of the output pulse depends on the laser mechanisms discussed below and will probably be in the range 50 ps to 1 ns for the new family of ASE lasers.

The fluorescent spectral line width of the laser is usually determined by Doppler broadening giving $\Delta\nu/\nu \sim 10^{-4}$, but laser spectral narrowing further reduces the line width in the laser beam by a factor of ~ 4 . The spectral brightness for the previously discussed parameters is plotted in Fig. 1, showing the dramatic increase relative to even laser produced plasma sources.

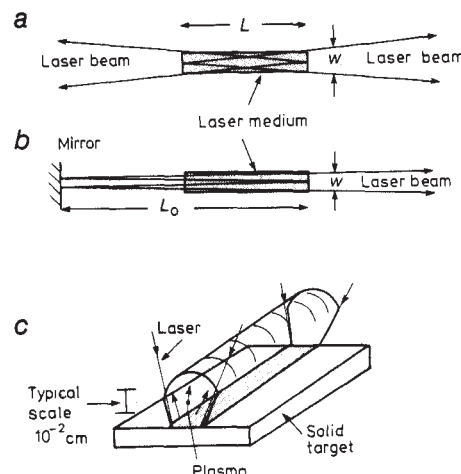


Fig. 3 Schematic single-pass (a) and double-pass (b) ASE lasers. c, Schematic model of plasma production by a laser beam focused on to a line.

Emission from an ASE laser is coherent in time only for a period of the order of the inverse spectral bandwidth $\Delta\nu^{-1}$ which from the earlier discussion means $\sim 4 \times 10^4$ cycles of the electromagnetic wave. The spatial coherence across the wavefront is determined by the rather large number ($n = w^2/L\lambda$) of transverse modes excited within the laser beam angle w/L . Transverse spatial coherence is restricted therefore to regions of dimensions of about $1/n$ (typically $1/50$) of the beam diameter. The use of a single mirror to double pass the laser medium offers improvements in spatial coherence through reducing the number of transverse modes.

Experiments and laser schemes

Proving population inversion and amplification. High-power pulsed lasers focused on to solid targets produce small hot expanding plasmas with dimensions 0.1 – 0.2 mm⁷. The existence of population inversion on XUV transitions was first determined and quantified in such plasmas from the ratio of intensities of spectral lines originating from the potential laser levels (for example, in Fig. 2 the population ratio n_3/n_2 is given by the intensity ratio I_{31}/I_{21} and the absolute intensity I_{31} gives the absolute number density of excited states n_3 ⁸. Measurements with resolution in space and time were developed for greater accuracy⁹.

Demonstration of population inversion did not prove the presence of net amplification because of the possibility of competing absorption processes such as the photoionization of other ions in the plasmas. In the next phase of experiments direct evidence of amplification was sought using elongated plasmas produced most often in a line focus of length $L = 1$ or 2 mm (this was limited by the available laser power) as illustrated in Fig. 3(c). GL values were measured from the ratio $R(L)$ of spectral line intensities for the laser transition along the axis and transverse to the axis of the plasma. This ratio is theoretically $(\exp(GL) - 1)/GL$ ¹⁰. Earlier experiments were generally not very accurate because GL values were small.

A few very recent experiments have used much longer line foci with $L = 1$ – 2 cm, and new time-resolving XUV spectrometers. With larger GL , variation of the length L between, say L_1 and L_2 allows a measurement of the ratio $I(L_1)/I(L_2)$ which is theoretically $(1 - \exp(GL_1))/(1 - \exp(GL_2))$. Direct evidence of exponential amplification has been obtained in this way. In particular the recent experiments at the LLNL showed GL values ≤ 6.5 (ref. 4).

3p–3s transitions in neon-like ions. The LLNL experiment used an idea which is derived from the mechanism of the familiar argon ion laser and was first suggested in the USSR¹¹. Other researchers followed up the concept^{12,13}, but most effort was confined to the USSR¹⁴ until a few years ago when, as part of

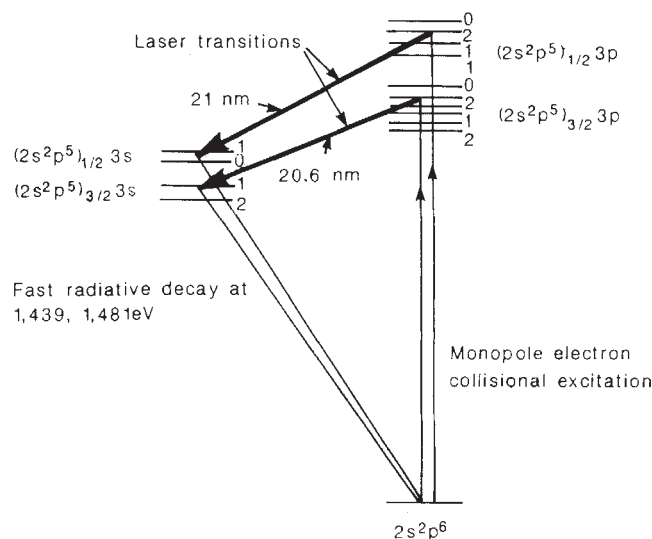


Fig. 4 Energy-level structure of Se^{24+} , after D. L. L. Matthews (personal communication).

a major campaign to test XUV laser options, the LLNL team began serious theoretical studies and experimental preparation for large-scale experiments using the two most powerful laser beams in the world from their new neodymium glass laser Novette.

The energy levels of the neon-like ions are illustrated in Fig. 4; the laser relies upon the fact that the ion has a closed shell of 2p electrons from which excitation by collision with plasma electrons populates the 3p states (as well as 3s and 3d states). There is no quantum mechanically allowed radiative decay from 3p to 2p, but there are strongly allowed transitions from 3p to 3s and 3s to 2p. If the thermal limit lies above the 3s level, then 3s to 2p radiative decay to the ground state reduces the 3s population and creates 3p to 3s population inversion.

The neon-like ion of selenium (Se^{24+}) has 3p to 3s transition energies of ~ 60 eV and was chosen for the LLNL experiments. A very thin (75 nm) layer of selenium was used to avoid reabsorption of 3s to 2p resonance radiation. The final important choice was to use a thin polymer film (50 nm) to support the selenium layer so that, when the target was irradiated with a 450 ps pulse at $5 \times 10^{13} \text{ W cm}^{-2}$, the whole target thickness was uniformly heated producing the Se^{24+} ions at a temperature of $\sim 5 \times 10^6$ K and causing an explosive expansion to a fairly uniform lower density (10^{21} electrons cm^{-3} or $3 \times 10^{-3} \text{ g cm}^{-3}$), at which maximum amplification occurred. The uniform density is vital to reduce refraction of the XUV laser beam as it travels along the 200 nm thick expanded layer of selenium plasma. Detailed calculations¹⁵ suggested amplification coefficients as high as $G = 10 \text{ cm}^{-1}$.

The two large (50 cm) beams of the neodymium glass laser were focused to a $12 \text{ mm} \times 200 \mu\text{m}$ line focus allowing up to 2 cm length to be irradiated. Sophisticated time resolving XUV spectrometers viewed the emission axially and transversely and Fig. 5 shows the clinching results. Two of the 3p to 3s lines at 20.6 and 21.0 nm were amplified by up to 500 times along the axis, indicating $G = 5.5 \text{ cm}^{-1}$. The output power was $> 100 \text{ W}$ with a divergence of 10^{-2} rad in a 300 ps pulse. A thermal source would need a temperature of $> 400,000,000 \text{ K}$ for this brightness. Despite the extreme brightness this is still only $\sim 1/3,000$ of the saturated output power which could be obtained with a modest twofold increase in GL . The importance of this result is considerable because it is the first absolutely incontrovertible evidence of high XUV gain. Limited isoelectronic scaling possibilities were also shown in the LLNL work with evidence of amplification in neon-like Yttrium at shorter wavelengths $\sim 15 \text{ nm}$.

Recombination. Population inversions are readily created in rapidly recombining plasmas, though this is not a mechanism

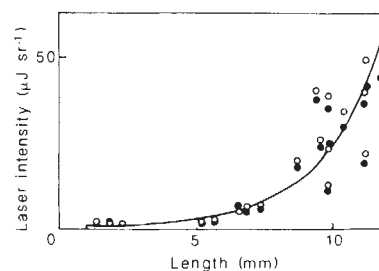


Fig. 5 Exponentially increasing output as a function of length demonstrating single-pass laser amplification of spontaneous emission in the Se^{24+} laser beam (redrawn from ref. 4). Solid circles, 209 Å; open circles, 206 Å; $-g = 5 \text{ cm}^{-1}$.

in any common type of laser. The idea, again Russian in origin¹⁶, dates from 1965 but isoelectronic scaling to the XUV was considered in the United States and the United Kingdom in the early 1970s (refs 12, 17, 18). Recombination from a plasma of bare nuclei to form hydrogenic ions, at densities such that the thermal limit is at $n = 3$ or higher, can result in inversion of population between $n = 3$ and 2, because of rapid radiative decay depopulating the $n = 2$ level. This occurs only if the ground state has an anomalously low population, a situation which can be produced by sufficiently rapid cooling such that the relatively slow process of recombination to the ground state lags behind the fast collisional processes across the small energy gaps populating levels above $n = 2$. Fortunately, such rapid cooling is produced almost automatically when a laser-produced plasma expands adiabatically into a vacuum from an irradiated target⁸.

Carbon plasmas have been extensively studied. Early evidence showed population inversion on the 18.2 nm $n = 3 - n = 2$ transition at low densities^{8,9}. Subsequent attempts to produce laser action used cylindrical expansion from 1 to 2 mm long laser-heated carbon fibres. The fibre size was restricted to a few microns in diameter so that the mass of the expanding plasma was small enough to avoid reabsorption of resonance radiation on the 2 to 1 transition. The material of the fibre is heated initially to $kT = 500 \text{ eV}$, at a density $n_e = 10^{21} \text{ cm}^{-3}$. Heating occurs during a 200 ps laser pulse through absorption of between 2 and 20 Joules per cm length of the fibre. Rapid expansion and cooling in the following nanosecond reduces the density to 10^{19} cm^{-3} and the temperature to 20 eV. Recombination into the ground state is minimal. Detailed calculations indicate amplification coefficients from 1 to 5 cm^{-1} of duration $< 1 \text{ ns}$ ¹⁹. The lower amplification is produced paradoxically by higher absorbed energy in a non-critical mode of operation involving total burn through of the fibre. High amplification is obtained with lower mass plasmas produced by partial burn through of the fibre, but demands very precise conditions which are difficult to produce uniformly along by the length.

Experiments have shown evidence of GL up to 4, deduced from axial to transverse intensity ratios, but with poor reproducibility in the high gain mode¹⁰. Experiments in the stable low-gain mode have so far had insufficient plasma length for direct measurement of amplification ($< 2 \text{ mm}$) but have shown good agreement between measurements of the plasma hydrodynamic characteristics and computer modelling¹⁹.

A rather different approach to this kind of laser has been used in other experiments²⁰. Plasma expanding from a carbon target irradiated by a 500 J, 10 GW pulse from a CO_2 laser has been confined to a narrow 1 mm diameter cylinder by an axial magnetic field and rapid cooling has been caused by thermal radiation. Evidence from axial to transverse intensity ratios suggests GL values of the order of 6. Variation of the plasma length is not possible and the feasibility of making a useful laser is difficult to determine. An amplification coefficient of 3 cm^{-1} has also been reported for thin polymer film targets similar to those used in the Se^{24+} laser and radiation cooling from the Se is thought to compensate for the reduced adiabatic cooling of a planar expansion²¹.

Recombination of helium-like (2-electron) ions to the lithium-like (3-electron) configuration is more complex to analyse but is the basis of an almost accidentally discovered mechanism for XUV laser action at 11 nm which has shown quite significant amplification ($G \sim 3 \text{ cm}^{-1}$)^{22,23}. The best results were obtained using 3 GW, 20 ns Nd glass laser pulses and an 8 mm line focus on solid aluminium targets. The laser transition is the 5f-to-3d transition of lithium-like Al^{10+} . The GL value has been determined from the intensity ratio for different plasma lengths as discussed earlier. Measurements have also included time-resolved spectroscopy showing the duration of the gain to be a few ns. Isoelectronic scaling has been demonstrated in similar experiments with magnesium. There are problems however with competing absorption processes which are not well understood. The recombination lasers have amplification coefficient comparable to the 3p–3s lasers and are more attractive for scaling to short wavelength because they use transitions with a change of principal quantum number with strong Z^2 scaling of the transition energy in isoelectronic sequences. They suffer less from refraction because the plasma density is much lower (10^{19} cm^{-3} compared with 10^{21} cm^{-3}) allowing longer plasmas and larger GL values.

Resonant photopumping. Selective excitation by photons is used in many optically pumped lasers and the method can in principle be extended into the XUV as first suggested by Russian authors in 1975 (ref. 24) and elaborated on in subsequent more detailed studies^{25,26}. Current thinking centres on using intense resonance line emission from one species of ion in a hot laser-produced plasma, with a brightness at the line centre corresponding to a black body with $kT_{\text{H}} \sim 200 \text{ eV}$, to resonantly excite another ion species in an adjacent plasma, for example, the $1s^2$ – $1s2p$ resonance line of Na^{9+} at 1.1 nm is absorbed by the $1s^2$ – $1s4p$ transition of Ne^{8+} . If the second plasma has a low temperature, $kT_{\text{c}} \ll kT_{\text{H}}$, and a density such that the thermal limit is above the laser levels, the $n=4$ level may be populated to produce population inversions, such as in the above example for the $4 \rightarrow 2$ transition at 5.8 nm.

Predicted amplification is as high as $G = 100 \text{ cm}^{-1}$ (ref. 26) and potential laser wavelength can be quite short as in the example given. Many such resonances have been identified in H, He, Li and F-like ions²⁵ but as yet there has been no success in demonstrating gain experimentally. But experimental work is at an early stage and the approach has considerable potential.

Other approaches. The free electron laser (FEL) mechanism in which coherent radiation is produced by deflection of a relativistic electron beam in a periodic magnetic field²⁷ is in theory scalable to any wavelength. But in practice, FEL devices have only been used to generate far infrared, infrared and visible emission, with progressively more difficult problems arising for operation at shorter wavelengths because of a $\lambda^{3/2}$ scaling of the amplification coefficient. It has been argued, however, that operation in the vacuum UV and XUV down to 50 nm might be achieved with magnetic wigglers²⁸ and that even X-ray operation at 0.5 nm could be contemplated with an electromagnetic wave as the 'optical wiggler'²⁹. Major and costly technical problems remain to be solved, including providing both an XUV resonator and a suitable storage ring or Linac of exceptionally high current, low electron-beam divergence and small electron-energy spread. The result could be a frequency tunable laser of high average power in a continuous train of pulses similar to a synchrotron radiation source. (Low power coherent radiation can be obtained from any source by frequency and angle selection and in the XUV synchrotron radiation sources equipped with undulators have been used for this purpose.)

Sudden transfer of stored excitation from long-lived to short-lived states can avoid the need for high fluorescent power and schemes based on charge exchange collisions³⁰ and laser excitation of metastable states^{31,32} are of interest. X-ray laser action at a photon energy of $\sim 1 \text{ keV}$ is reported to have been obtained using a thermonuclear explosion as the pumping source³³. No details of this work have been published, though it may be surmised that the mechanism would involve photon pumping.

The device is clearly not a laboratory laser and its potential applications are in the area of the United States' 'Star Wars' programme.

Future applications of XUV lasers

If, as expected, laboratory XUV lasers are developed successfully in the next few years they will offer powerful, MW to GW single pulses at photon energies (or wavelengths) from $\sim 50 \text{ eV}$ (25 nm) to perhaps 500 eV (2.5 nm). These lasers will only be available at major laboratories where there are exceptionally powerful optical frequency lasers to create the plasmas. So the likely applications will be scientific rather than technological, with powerful XUV lasers complementing other sources, notably synchrotron radiation and thermal X-rays from laser-produced plasmas. The XUV lasers will be used to best effect where time resolution, intensity, spectral purity and coherence are important. Speculation on future applications is likely to be inaccurate since unforeseen developments are usually most significant. The most probable initial applications will be extrapolations of current lines of research:

Soft X-ray microscopy in biology. Recent work with single-pulse soft X-ray emission from laser-produced plasmas has demonstrated high-resolution (50 nm) images of living cells using contact radiography on soft X-ray resists³. A laser source offers the possibility of holographic imaging with better resolution and three-dimensional information, with important implications for progress in biology.

Diffraction from crystals. Analysis of molecular structures from single crystal Laue diffraction is fundamental to biophysics², but suffers from the lack of any information on the phase of the diffraction peaks. A coherent laser source could, by holographic methods, give extra information, leading to unambiguous structure determination. The problem is that wavelengths an order of magnitude less than those envisaged here are needed to resolve the molecular structures, so although some low resolution work of this kind could be contemplated, the major prizes would remain inaccessible. Single-pulse diffraction measurements on transient phenomena³ would be possible, for example, thermal and chemical transformations in thin films and surfaces—but only on systems where crystal structures were larger than the wavelength.

Material processing. Research into holographic replication of submicron patterns in lithography for application to microelectronics, integrated optics or diffraction grating device fabrication could achieve still smaller scale using powerful XUV laser sources.

Atomic physics. Sub-nanosecond selective excitation or ionization of atomic or ionic beams with an XUV laser would open up new possibilities in the high-resolution spectroscopic study of atoms, ions and molecules and would be particularly appropriate for XUV lasers in the low-energy region ($h\nu \leq 50 \text{ eV}$) with the high laser power being important for adequate single-pulse yields and the monochromaticity for high spectral resolution.

Laser physics and nonlinear optics. There is considerable interest in improving the characteristics of the new lasers, developing resonators, schemes for shorter and more powerful pulse generation (perhaps by injecting harmonically generated picosecond pulses into plasma amplifiers) and exploring nonlinear optical effects for the first time at XUV frequencies.

- Waynant, R. W. & Elton, R. C. *Proc. IEEE* **14**, 1059–1092 (1976).
- Synchrotron Radiation Research* (eds Winick, M. & Doniach, S.) (Plenum, 1980).
- Key, M. H. *Plasma Phys. and Contr. Fusion Res.* **26**, 1383–1394 (1984).
- Mathews, D. L. *et al. Phys. Rev. Lett.* **34**, 110–113 (1985).
- Reintjes, J. *et al. Phys. Rev. Lett.* **30**, 480–482 (1977).
- Bates, D. R., Kingston, A. E. and McWhirter, R. W. P. *Proc. R. Soc. A* **267**, 297–312; **A270**, 155–167 (1962).
- Key, M. H. & Hutcheon, R. J. *Rep. Prog. Phys.* **16**, 201–280 (1980).
- Irons, F. E. & Peacock, N. J. *J. Phys.* **B7**, 1109–1112 (1974).
- Key, M. H., Lewis, C. L. S. & Lamb, M. J. *Opt. Comm.* **28**, 331–335 (1979).
- Jacoby, D., Pert, G. J., Shorrocks, L. D. & Tallents, G. J. *J. Phys.* **B15**, 3557–3580 (1982).
- Molchanov, A. G. *Sov. Phys. Uspekhi* **15**, 124–129 (1972).
- Key, M. H. *Gordon Res. Conf. Laser Interaction with Matter* (1973).
- Elton, R. C. *Appl. Opt.* **14**, 97–101 (1975).

14. Vinogradov, A. V. & Shiyaptiev, V. M. *Sov. J. quant. Electr.* **13**, 303-306 (1983).
15. Rosen, M. D. *et al. Phys. Rev. Lett.* **54**, 106-109 (1985).
16. Gudzenko, L. I. & Shelepin, L. N. *Sov. Phys. Dokl.* **10**, 147-149 (1965).
17. Slutz, S., Zimmerman, G., Lokke, W., Chapline, G. & Wood, L. *Bull. Am. Phys. Soc.* **17**, 972 (1972).
18. Pert, G. J. *1st Natn. QE Conf.* (Manchester 1973).
19. Pert, G. J. *et al. in Laser Techniques in the Extreme UV* (eds Harries, S. E. & Lucatorto, T. B.) (AIP, 1984).
20. Suckewer, S., Keane, C., Milchberg, M., Skinner, C. H. & Voorhees, D. *in Laser Techniques in the Extreme UV* (eds Harris, S. E. & Lucatorto, T. B.) (AIP, 1984).
21. Seely, J. P. *et al. Centre for Space Research, Preprint* (1985).
22. Jaegle, P. *et al. J. Phys. B* (in the press).
23. Jaegle, P. *et al. in Laser Techniques in the Extreme UV* (AIP, 1984).
24. Vinogradov, A. V., Sobelman, I. I. & Yukov, E. A. *Sov. J. quant. Electr.* **5**, 59-63 (1975).
25. Dixon, R. H. & Elton, R. C. *J. Opt. Soc. Am.* **2**, 232-238 (1984).
26. Apruzeze, J. P., Davis, J. & Witney, K. G. *JAP* **53**, 4020-4027 (1982).
27. *IEEE J. Quant. Electr. Spec. Issue on Free Electron Lasers QE 19* (1983).
28. Goldstein, J. C., Newman, B. E., Cooper, R. K. & Comley, J. C. *in Laser Techniques in the Extreme UV* (eds Harris, S. E. & Lucatorto, T. R.) (AIP, 1984).
29. Dobiasch, P., Megstre, P. & Scully, M. O. *IEEE J. quant. Electr. QE 19* 1812-1820 (1983).
30. Dixon, R. H., Seely, J. F. & Elton, R. C. *Phys. Rev. Lett.* **40**, 122-125 (1978).
31. Harris, S. E. *Opt. Lett.* **5**, 1-3 (1980).
32. Wong, J. C., Caro, R. G. & Harries, S. E. *Phys. Rev. Lett.* **51**, 767-770 (1983).
33. Robinson, C. A. *Aviation Week and Space Technology* 23 February, 25-27 (1981).

Lasers, nonlinear optics and optical computers

S. D. Smith

Department of Physics, Heriot-Watt University, Edinburgh EH14 4AS, UK

Devices whose optical properties change with light intensity have opened the path towards the optical computer. Optical switches and signals can now undertake all binary logic operations and complete optical digital circuits have been constructed.

FROM the earliest days of the laser, the large intensities available in laser beams permitted irradiances of the order of MW cm^{-2} incident on material by simple focusing. With such fluxes the amplitude of the oscillating electric field, E , readily becomes comparable to interatomic fields such that the optical properties of the medium vary with the passage of the wave. The polarization, P , induced by the radiation in the simplest case (of a single monochromatic wave propagating through a crystal of high symmetry) has the form

$$P = \chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 \dots \quad (1)$$

The susceptibilities $\chi^{(1)}$, $\chi^{(2)}$, $\chi^{(3)}$... are related to the optical properties of the medium: as examples, the linear optical properties are related to $\chi^{(1)}$, the second harmonic generation is determined by the magnitude of $\chi^{(2)}$ and the intensity dependence of the refractive index is related to $\chi^{(3)}$. The science of nonlinear optics developed with the observation of second harmonic generation by Franken *et al.*¹ in 1961 and the theoretical prediction of a series of nonlinear effects by Bloembergen² and co-workers.

The large powers required in the early work made it seem unlikely that digital optical computing, using optical circuitry with nonlinear optical circuit elements, would ever be practical. It was nearly 20 years before the experimental discovery that a 30-mW continuous wave (CW) laser beam could be significantly self-defocused in a narrow gap semiconductor produced the surprising conclusion that third order nonlinearities ($\chi^{(3)}$) could have a value of the order of 10^9 times larger than the nonlinearities envisaged by Franken and Bloembergen^{1,2}.

To understand the significance of this giant nonlinearity to a possible future of optical computing, consider the physics of computation (see ref. 3). Information is essentially stored as energy and switching a device from logic 0 level to logic 1 level requires a definite switching energy. This energy must be greater than the thermal energy of the device kT (and will usually be several hundred times kT) and scales as the size of the device. There will usually also be a trade-off between speed and power. A further absolute requirement for electronic or optical logic is the necessity of restoring a logic level after each switching action so that errors resulting from imperfect devices and signals do not accumulate. Restoring logic must have power gain and this power is normally drawn from a power supply independent of the signal channel. (The requirement of power gain is obviously necessary, that is, one switching device must be capable of

driving at least the next in the series so that a circuit containing multiple elements can be constructed.) These considerations appear to be applicable to all systems of signal processing and computing.

The response of semiconductor micro-electronics to the high data rate requirements of digital signal processing and computing has been to increase switching speeds and further miniaturize components in the form of very large-scale integration (VLSI). Transistors made from gallium arsenide have been reported with effective switching times as fast as 12 ps. However, this will not necessarily solve the problem of coping with high data rate as the processing time in conventional computers is many times the logic switching time because of the necessity of transferring information to the next part of the circuit. This involves capacitance time constant limits as pointed out, for example, in ref. 3; VLSI does not solve the time-constant problem because, as the length of a wire shrinks by a factor α and the cross-sectional area of the wire is reduced by a factor α^2 , the capacitance C of the wire decreases by this factor α while the resistance R increases by the same amount. Thus, the time constant RC remains the same and the input charging time remains unaltered independent of scaling.

Turning to the computer as a whole, the standard method of communication in use today connects the logic unit with the memory through an address device. This reduces the number of interconnections but can only address one storage element at a time. This widely used scheme was first suggested by von Neumann but, rather than being given credit for this most practical innovation, he is now rather undeservedly blamed for this so-called "von Neumann bottleneck". The timing problems associated with circulating logic signals around a one-dimensional processor of this type ('clock skew') combine to indicate that future problems in digital computers are likely to be those of communication. This may apply at all levels (architectural, bus and chip) and stems from the use of time multiplex to compensate for the inability of electrical methods to communicate many channels of information in parallel.

Digital optics

Present practice has seen the invasion of electronically based communication by optical methods through the use of optical fibres in long-range telephone lines. The higher carrier frequency used gives potentially higher bandwidth, although electronic

limitations on modulation techniques have restricted our ability to exploit this greater information carrying capacity fully. Currently, long-range transmission made possible by low signal attenuation has been used. The use of optics for processing information has so far been handicapped by the absence of optical circuit elements. The optical methods have promising implications in areas that are currently difficult for existing technologies; these include image processing and recognition, sorting, radar array signal processing, machine vision and artificial intelligence.

During the past 20 years there has also been a body of work often called optical computing practised by optical researchers using linear processing devices, for example, the use of spatial filters and Fourier transform processes in image processing. Ideas include the use of symbolic substitution, residue arithmetic, vector-matrix multiplication and primitive processing experiments using a liquid crystal light valve as a hybrid (that is, optics with electronics) spatial light modulator but not itself capable of all-optical logic action. Most of the reports are theoretical proposals of what might be done (see ref. 4).

The major stumbling block in the development of digital optics has been the absence of nonlinear all-optical circuit elements, preferably capable of fabrication in the form of two-dimensional arrays and of small enough size to have small switching energies and high speeds.

The main purpose of this article is to review the latest state of research in all-optical nonlinear logic switches, amplifiers and memories and to show that, with the experimental realization of the first all-optical circuits, the time is now ripe for the combination of these two previously separate research efforts, optical bistability and linear optical computing techniques.

All-optical circuit elements

Recent progress, which has produced practical all-optical circuit elements, is based on the combination of optical nonlinearity and feedback. This has led to the concept of 'optical bistability' and hence to a whole family of devices based on a common set of physical and mathematical principles. The series of devices includes optical logic gates, bistable memories, amplifiers (sometimes termed optical transistors or transphasors) and power limiters (see ref. 5 for a recent Royal Society, London, conference). The meeting of the International Commission of Optics in Sapporo⁶ presented six papers on optical computing and began the process of introducing the optical bistability community to the optical information processors. The recent Optical Society of America conference⁷ at Lake Tahoe continued this process; a three-element loop processor, using all-optical interactions, was reported, together with experiments on parallel optical computing, using liquid crystal light valves as the nonlinear elements.

A seminal paper on optically bistable devices was that of Szöke *et al.*⁸ who in 1969 proposed that a Fabry-Perot optical resonator containing a saturable absorber as its spacer layer could exhibit two states of transmission for the same input intensity. This simple bistable action was predicted to arise from the existence of a high internal optical field at constructive interference given that sufficient intensity had been incident on the resonator to bleach the absorber. To reach this condition required a greater input intensity than that required to maintain it. By contrast, at low input intensity the non-bleached absorption held the transmission of the device at a low level. In practice this condition is quite hard to achieve experimentally and the experiments described do not in fact show optical bistability. Observation of optical bistability was not made until 1976 when Gibbs *et al.*⁹, using an interferometer containing sodium vapour, observed bistable transmission but deduced that the dominant mechanism was refractive, involving a shift in resonator frequency, rather than absorptive. Nevertheless, effective refractive nonlinearity resulted from a saturation of the atomic absorption. Such a device, although using only milliwatts of power, was relatively large (centimetres in length) and relatively slow (milliseconds) compared with electronic circuit components.

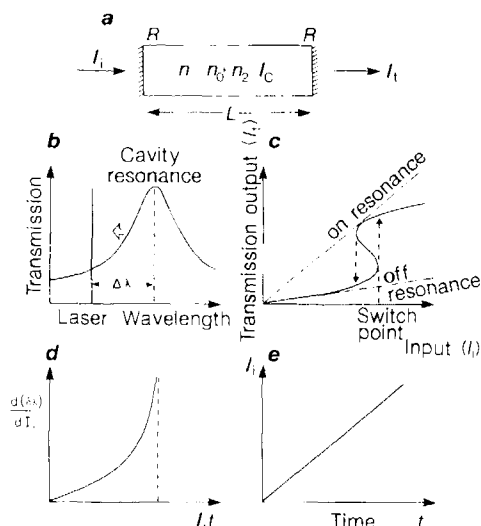


Fig. 1 The dynamics of switching in a Fabry-Perot etalon (see text for explanation of individual diagrams).

In the same year, my colleagues and I made the surprising discovery of giant nonlinear refraction, defined as a function of the intensity, I , from

$$n = n_0 + n_2 I \quad (2)$$

(where the nonlinear refractive index n_2 can be measured in units of cm^2 per kW) present in the narrow bandgap semiconductor InSb. The immediate implication was that a bistable resonator could be made of micrometre dimension and, as the effect was shown to be electronic, would be fast, probably on a nanosecond timescale. A second implication was that one beam could modulate the optical properties of a small slice of semiconductor and affect a second beam, thus making an optical modulator or optical transistor. The details of both the refractive and associated absorptive nonlinearities at milliwatt powers are described in ref. 10 but, retrospectively, the effects could be recognized in the earlier work on the spin-flip Raman laser. Both the device possibilities described above had been practically realized in InSb by 1979 (ref. 11) in which continuous wave laser beams were used, leading to steady-state operation and true optical bistability, as well as the observation of gain in an optical transistor¹².

Simultaneously and independently, optical bistability was reported by Gibbs *et al.*¹³ using pulsed dye laser radiation in the larger gap semiconductor GaAs. Larger absorption and a smaller nonlinearity in this material, however, prevented steady-state operation and this observation was therefore quasi-dynamic and did not permit the demonstration of logic levels or differential amplification.

Origin of giant nonlinearities

The physical explanation of the large nonlinearities in both these semiconductors involves the excitation of electrons to give some degree of saturation or 'blocking'. In the case of InSb, exciton effects are negligible in the conditions of the experiment and a plausible explanation has been given by Miller *et al.*¹⁴ in terms of a 'dynamic Burstein-Moss' shift of the band edge. Physically, a number of electrons ($\sim 10^{15}$ – 10^{16} cm^{-3}) are excited into lower conduction states by the laser photons. The incident intensity I generates an equilibrium number of electrons. Subsequent band filling, following thermalization, modifies the absorption edge and by application of the Kramers-Kronig relationship causes a change in refractive index,

$$\Delta n = \frac{\hbar c}{\pi} \int_0^\infty \frac{\Delta \alpha(\hbar \omega')}{(\hbar \omega')^2 - (\hbar \omega)^2} d(\hbar \omega') \quad (3)$$

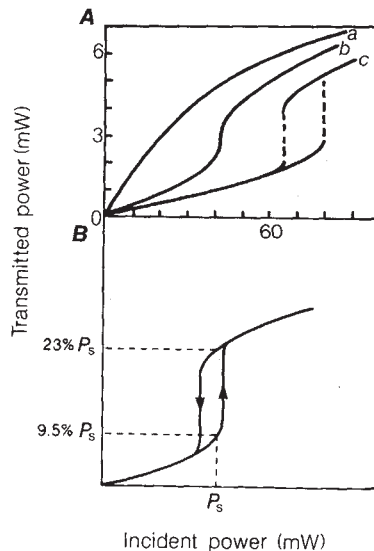


Fig. 2 **A**, Experimental observations of a family of characteristics of an interference filter bistable device illuminated by the argon ion laser line of wavelength 514 nm, obtained by changing the initial detuning from resonance of the etalon. The family of curves *a-c* shows the characteristics obtained for different values of the initial detuning parameter $\Delta\lambda$ set by the angle of incidence. If the initial condition is close to resonance, power limiting action is obtained (*a*) as the transmittance can only fall as intensity rises. As $\Delta\lambda$ is increased, curve *b* begins to kink, showing a greater change in output than for an incremental change in input, and thus exhibiting differential gain. This is responsible for optical transistor or transphaser action. Further increase of $\Delta\lambda$ leads to bistable loops of varying width (*c*). **B**, Experimental observation of optical bistability in the interference filter illuminated by light of wavelength 528 nm. The stability of the system is sufficient to allow operation to within $\sim 4\%$ of the power required to switch.

where $\alpha(\hbar\omega')$ is the interband absorption coefficient at photon energy $\hbar\omega'$. With $\Delta n/\Delta I \approx n_2$ we can obtain an analytical expression for n_2 . This calculation gives a good description of the resonant refractive nonlinearity in InSb, InAs and CdHgTe.

Device physics

The simplest configuration which provides optical feedback is a simple Fabry-Perot etalon containing a nonlinear refractive material (Fig. 1a). Its optical thickness is given by

$$nL = (n_0 + n_2 I_c) L \quad (4)$$

(where L is the thickness of the material) and this changes with the internal intensity I_c . Consider now the transmission of such an interferometer as a function of wavelength (Fig. 1b): if we start in an initial condition where the illuminating laser wavelength is detuned from maximum transmission by a wavelength increment $\Delta\lambda$ (Fig. 1b), we see from Fig. 1c that the relation between output and input would give rise to, in the case of a linear device, a straight line of low slope: if the device were tuned to resonance and there were no absorption loss, output would be related to input by a line at 45° . If we now increase the intensity from the initial condition, the nonlinear resonator tends towards resonance as its optical thickness changes with intensity. This gives rise to a nonlinear relation between output and input. However, as we approach resonance, the internal field circulating within the resonator itself builds up according to

$$I_c = I_i T(\lambda) (1 + R) / (1 - R) \quad (5)$$

where I_i is the incident intensity, $T(\lambda)$ is the transmission as a function of wavelength (as in Fig. 1b) and R is the (constant) reflectivity of the resonator mirrors. Thus, at resonance the internal intensity is at its maximum where $T = 1$ and is amplified by the term $(1 + R)/(1 - R)$. This gives rise to positive feedback. As resonance is approached, the internal field builds up and

the rate of approach to resonance depends on the change in optical thickness which itself depends on the magnitude of the internal field. The rate of approach to resonance thus speeds up. This can be readily expressed through the expression

$$\frac{d(\Delta\lambda)}{dI_i} = T(\lambda) / \left(\frac{T_{\max}}{2n_2 L} - I_i \frac{dT}{d\lambda} \right) \quad (6)$$

Figure 1d shows a plot of the rate of approach to resonance as a function of either incident intensity I_i (or as a function of time if we assume a linear ramp of I_i against time; Fig. 1e). As I_i increases, $dT/d\lambda$ also varies, the denominator of equation (6) tends to zero and the rate of approach to resonance becomes infinite, leading to rapid switching. This gives a physical feel for the dynamics of the switching process and leads to optical bistability (Fig. 1c), that is, two values of the transmission for one of the intensity. Two features are noteworthy: (1) a large value of finesse (ratio of spacing to half-width of the transmission peaks) makes it easier to obtain bistability; and (2) the form of the characteristic (output versus input) changes with the initial conditions, that is, it depends on $\Delta\lambda$.

Such a set of characteristics is illustrated in Fig. 2, giving a set of experimental results for nonlinear interference filters incorporating ZnSe films as the active layers (curves *a-c*). Thus, by a simple change of initial conditions a whole family of optical devices can be produced.

Figure of merit for optical devices

An analysis of the factors concerned in the design of such optical circuit elements has been given by Miller¹⁵, who shows that the quantity which gives a figure of merit (in terms of nonlinear Fabry-Perot cavity optical switching in the presence of linear absorption, α) is $n_2/\lambda\alpha$. This determines the lowest critical value of input irradiance I_s for a device of given size to obtain bistable switching or nonlinear characteristic:

$$I_s = \frac{\alpha\lambda}{n_2} f(R, \alpha L) \quad (7)$$

The result is physically sensible as the switching power will be lower for a larger nonlinearity n_2 , the shorter the wavelength the smaller will be the refractive change required to effect a change from constructive to destructive interference (that is, $\Delta(nL) = \lambda/2$) and for a smaller absorption (assumed linear in this analysis) the longer the device may be for a given loss. The function f optimizes cavity properties. Values of $n_2 = 0.1 \text{ cm}^2 \text{ kW}^{-1}$ and $\alpha = 10 \text{ cm}^{-1}$ give useful devices of thickness (L) 50–200 μm , bearing in mind that a fractional change $\Delta n/n \sim 10^{-3}$ is usually required.

To obtain favourable values for $n_2/\lambda\alpha$, we make particular use of effects resonant with the energy gap E_G in semiconducting materials. For electronic nonlinearities it can be shown that $n_2 \sim 1/E_G^3$ so that $I_s \sim 1/\lambda^2$. In addition, the considerations made earlier suggest that the smallest possible devices should be constructed. Diffraction limits suggest that the area limit will be $\sim (\lambda/n)^2$ and thus, although the nonlinearity is clearly large at longer wavelengths for small gap materials, the interference conditions and device size favour shorter wavelengths. Analysis of the detail of the frequency dependence of the nonlinearity, together with the attendant unwanted losses, has not so far indicated an optimum wavelength.

Semiconducting compounds have shown promising results: InSb with a typical working wavelength (5.5 μm) corresponding to $1,820 \text{ cm}^{-1}$ is one material for which there are sufficient available laser frequencies to undertake a detailed examination of the frequency and hence resonant behaviour of n_2 and α near the band gap^{5,14}.

GaAs has also been investigated^{5,13} but differs by having a strong exciton feature near the absorption edge. The nonlinearity $n_2 \sim 10^{-3} \text{ cm}^2 \text{ kW}^{-1}$ is quite practicable for devices, but the absorption coefficient in epitaxially grown material is so far giving values of $\alpha \sim 10^4 \text{ cm}^{-1}$. Thus, thicknesses are restricted to a few micrometres and thermal stability poses a problem.

Both the above materials show negative values of n_2 caused by electronic effects. Carrier lifetimes are in the range of tens to hundreds of nanoseconds. The nonlinearities can be 'switched on' more quickly than this time interval by rapidly introducing carriers with relatively intense pulses. The relative figures of merit between GaAs and InSb favour InSb by a factor of 10^5 , which has given InSb the advantage of steady-state operation but the disadvantage of a low operating temperature of 77 K, whereas GaAs has operated at room temperature and at shorter wavelengths. So far there are no reports of true steady-state operation for GaAs material. Other materials, which have been used in purely optical switching devices, are reported in ref. 5.

There is a second useful form of nonlinearity involving thermal shift of the band edge resulting from temperature rise of the bulk of the nonlinear material. This effect also resonates with the band edge and is associated with moderate values of absorption coefficient. The quantity taking the place of n_2 is an effective $n_2 = (dn/dT)\alpha(LL'/\kappa_s)$ where L , L' introduce dimensions of film and substrate thickness and κ_s is the thermal conductivity of the substrate. Interference thin film structures using ZnSe (ref. 16) have given promising results for refractive switching, and bulk ZnSe, CdS and GaAs also show 'optical bistability by increasing absorption' where the thermal shift introduces the required feedback⁵.

Requirements for optical computing

Since our proposition is to undertake logic operation by means of these optical circuit elements, we can define some of the requirements if they are to be put together in the form of optical circuits to construct an optical computer. Requirements are as follows: (1) High contrast. A logic device needs to show a large change between logic 0 and logic 1 levels. (2) Steady-state bias. To make various different logic gates it is necessary to alter controllably optical bias levels. In terms of optical bistability this means that the device can be 'held' indefinitely at any point on the characteristic with a CW laser beam—the 'holding beam'—and implies a degree of thermal stability. Devices based on InSb at 77 K and on ZnSe at 300 K have been the first to show this behaviour. (3) External address. For logic functions it is clearly necessary that separate external signal beams can be combined with the holding beam to switch the device. The switching energy in fact is derived from the holding beam and this switched beam propagates in transmission or reflection as the output beam to the next devices in an optical circuit. (4) The elements must be cascable. This means that the output of one device must be sufficient to switch at least one succeeding device. The ability to set a CW holding beam near to switch point in fact fulfils this condition because the extra increment is then small compared with the change in output even in the presence of loss. As each device has its own 'power supply' (that is, holding beam), logic levels are restored. (5) Fan-out and fan-in. The probable advantages of parallel processing in optical devices emphasize the ability that one device can drive many succeeding devices, probably using free space propagation for address. The summed effect of several elements can readily be focused on one device to achieve fan-in. (6) High gain. Items (4) and (5) demand that the elements show a value of differential gain ratio of change in output to change in input of >1 . (7) Arrays. The technology shall ideally be such that two-dimensional arrays are easily constructed. (8) Power and speed. Low power per device is a necessity and this should preferably be of the order of milliwatts or less; this is aided by fabricating small-area devices. Speed and power will be interchangeable but speed itself will vary with use: for parallel arrays, micro-seconds will suffice and for one-dimensional circuits, sub-nanosecond or picosecond switching times are desirable.

Experimental results

Figure 3a shows an early result for an InSb resonator operating in transmission. The critical switching power is 20–25 mW incident on a diameter of $\sim 200 \mu\text{m}$. If the device is set up in

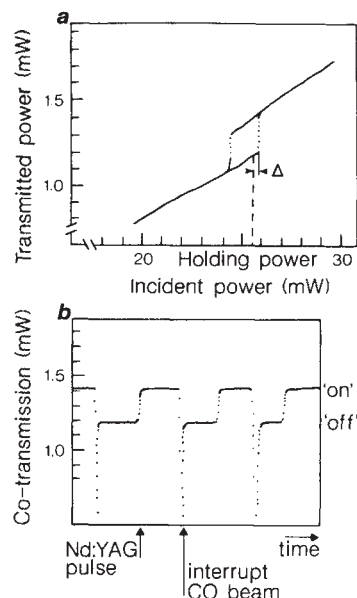


Fig. 3 *a*, Transmission characteristic of an InSb etalon illuminated by a steady-state carbon monoxide laser operating at a wavelength of $5 \mu\text{m}$. The laser power is constant at a level slightly below that required to switch. *b*, The change in transmission induced by the absorption of 35 ps long pulses from a Nd:YAG laser ($1.06 \mu\text{m}$) and total energy 5 nJ.

this way with the holding beam adjusted in intensity to be just short of the switch point, the device may be externally addressed. Figure 3b shows the results obtained using a single 35-ps long pulse of 5 nJ energy from a Nd:YAG laser. The arrival of the single pulse is sufficient to trigger the switching and the device remains in the 'on' state. Interrupting the holding beam returns the switch to the 'off' state, and as can be seen the logic levels are extremely stable. The device was further developed to show explicit AND gate operation by dividing the switching pulse with a beam splitter and observing that the switching energy was quite definite so that the device could be set to switch only when both pulses were incident (in addition to the steady-state holding beam) and would not switch with a single pulse. The device is therefore acting as an AND gate and an optical memory.

We may deduce that sufficient free carriers were induced near the surface of the InSb resonator to cause sufficient change in optical thickness to initiate switching within ~ 3 ps. W. Kaiser (personal communication) has shown recently that this nonlinearity can be switched 'on' in <7 ps. Therefore, the probable switching time is limited only by the macroscopic effect of internal field build-up within the resonator. This depends only on the round-trip time for the $210\text{-}\mu\text{m}$ thick etalon and amounts to ~ 8 ps. We may infer that picosecond timescale optical logic has been demonstrated.

First digital optical circuits

With CW steady-state hold, an InSb logic device can be switched with incremental power as little as $3 \mu\text{W}$ (Fig. 3). Thus, the output change of the device operating in reflection (Fig. 4), ~ 4 mW, should be sufficient to switch a succeeding device. A first experiment using two optical switches, A and B, is illustrated in Fig. 4. This first example of an optical logic circuit is equivalent to an XNOR gate.

Thin-film interference structures comprising alternate layers of high- and low-index material deposited on a flat substrate such as float glass are a convenient way of constructing a resonator equivalent to the devices described earlier. If sufficient nonlinearity can be induced in the very thin interference layers the technology has many advantages, particularly that very uniform devices covering many square centimetres can be easily made. Thus, large-area arrays are readily possible. Typical total thicknesses of films for a device with 15 layers are only $\sim 2\text{--}3 \mu\text{m}$.

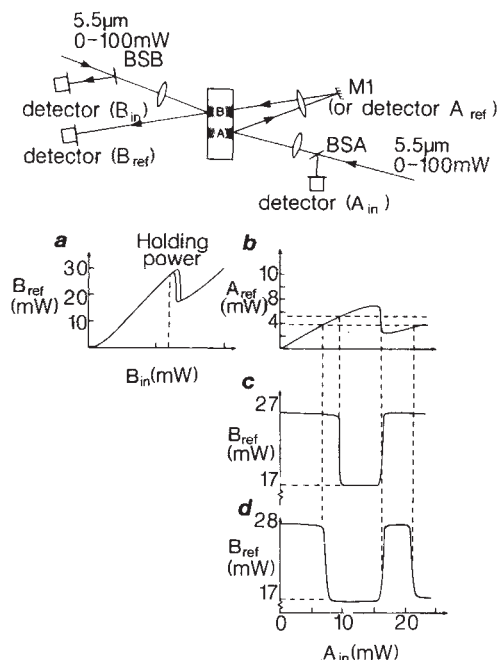


Fig. 4 The experimental arrangement used in the fabrication of an XNOR gate by the coupling of two InSb logic gates and the characteristics of the individual (a, b) and coupled (c, d) gates. Optical switch A is addressed by a holding beam which is reflected from the element and focused on to a second element, B, 500 μm away on the same crystal slice. This second element is held near to switch point by a second holding beam illuminating it from the opposite side. The output of A then acts as the address beam for B and the optical circuit output is observed as the reflection from B (B_{ref}). The figure also depicts the circuit inputs and outputs as the input to A (A_{in}) is steadily increased until a point is reached where its reflected output acting as address on B causes this second gate to switch to a lower state in reflection (c), (d). B remains in this lower state until A itself switches down in reflection, when B simultaneously switches up in reflection. B remains in the up state as the input to A is further increased until it again switches down when the reflected input from A is high enough. If A were to be held close to switch point and addressed with a pulse, information would be stored in A. If the holding beam on B is at the same time too far from switch point, the information will not be transferred from B. If this holding beam is then programmed to come within the range where B can be switched, the information can be transferred to B and by an appropriate reduction of the holding beam on B the information can be moved from A. The device can therefore be made to act as a shift register. With further holding beams defining optical circuit elements across the crystal, these forms of optical logic become indefinitely extensible.

The heat capacity of pixels of micrometre dimension is low and implies that incident power of ~ 10 mW can cause a temperature rise of 50°C in ~ 1 ns. The important physics lies in the control of heat sinking by means of the conductivity of a relatively massive substrate. Figure 2 shows the very promising results obtained using ZnSe as the active film which has a band gap conveniently resonant with argon ion laser wavelengths at 514 and 528 nm. The important property recently found is that these logic switches are sufficiently stable to permit construction of optical circuits of a similar nature to those already demonstrated with InSb^{5,17}.

Optical computers

I have described the contemporary performance of several optical circuit elements and shown that, in principle, all logic functions familiar in electronics can be reproduced by these nonlinear optical devices. Therefore, it should be possible to construct an all-optical computer.

There exists a considerable body of literature reporting the research of a number of groups on the subject of optical comput-

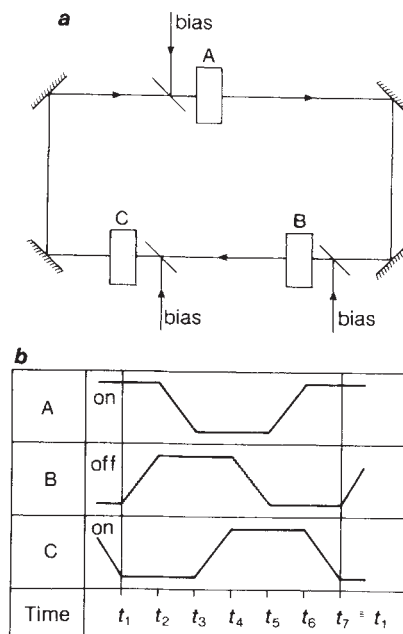


Fig. 5 a, A schematic model of the arrangement used to demonstrate a three-gate optical delay which forms the basis of a loop processor. b, The clocking of the individual bias levels to the three gates.

ing⁴. To date this has not included the use of the nonlinear elements discussed here. Nevertheless, the advantages of using optical methods for various computations have been apparent for several years. Historically, most optical computing systems have been of analog type in which information is stored and processed as a continuum of signal levels. There are major drawbacks to such analog systems, including limited flexibility, noise accumulation and input/output device limitations. The advantages of optical binary logic, now in principle possible with the devices described here, may be crucial in progressing optical computing into a new practical technology. One would wish to avoid as many photon-electron or electron-photon conversions as possible. The considerations can be made at several levels, including individual optical logic devices (gates and arrays of devices) as the first level. The second level considers optical communications interconnections and input/output among the logic gates, among arrays and among circuit boards or processors. The third level considers the possibilities of new computer architecture to take advantage of the inherent parallelism of optics.

Thus, one might see the advantages of optical methods in two separate ways: (1) Speed of switching and communication. I have shown that individual optical logic operations can almost certainly be performed on a picosecond timescale. Combining this with the output characteristics of CW mode-locked lasers, it is possible to conceive bit rates approaching frequencies of THz for an exact analogue of electronic digital logic. It is not yet clear that recovery times can be imposed on the optical logic switches at acceptable power levels to ensure continuous operation. The possibility exists of series-to-parallel and parallel-to-series conversion on picosecond timescales. Fast, high bandwidth communication will also be consistently available. Optical methods may defeat clock skew. (2) Parallelism and communication. The logic devices described here are at this time the only ones which have been capable of providing steady-state information holding and cascable operation. They have time constants in the range 100 ns–100 μs . This suggests that the first approach to the use of optics should emphasize the use of parallelism and accept cycle rates similar to existing electronic speeds. The intriguing intellectual challenge is to use the flexibility of optics to provide appropriate interconnections:

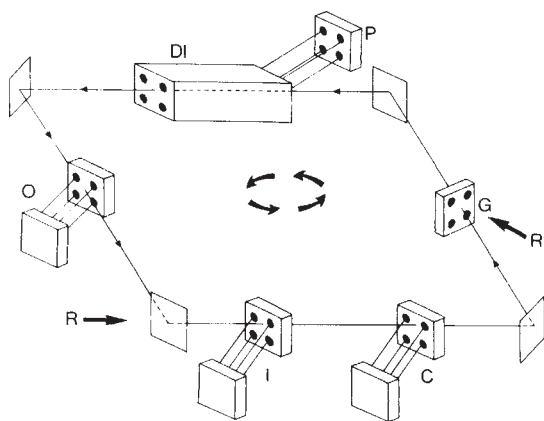


Fig. 6 The types of optical computer being proposed use the parallelism of optics by using arrays of elements. The iteration present in the type of problem which could be solved on such a computer is achieved through the use of a loop processor architecture. The individual arrays of elements may have to be flexible in their uses in the processor, identical arrays may be used as input (I), clock (C), gain (G), program (P), dynamic interconnect (DI) and output (O). Each array may require a separate power supply refresh (R).

requirements could be to shift a logic array pattern by one element on each cycle or to achieve a 'perfect shuffle'. Other schemes include a vector-matrix multiplication involving a fan-in from a matrix array to a column followed by rotation of this column between output and input for successive cycles.

In my laboratory a number of optical-computer architectural components have been designed including memory units, clocks, a programmable processing stage and a simple full adder using simultaneously the transmission and reflection from an optical gate¹⁸. The necessity to store the calculation before communicating to the next element is important in the optical case: propagation at the speed of light otherwise would mean that all elements would be simultaneously addressed (admittedly, avoiding clock skew³). Delays can be readily implemented by suitably programming the holding beams so that the optical bias is only set to receive a signal when the device is clear of other signals. As a first step, consistent with the elements described, this could be achieved electronically making an interesting interface to existing computer technology. At the time of writing, a three-element loop processor with optical bias delay loop clocking has been implemented. Acousto-optic modulators, controlled by a micro-computer, provide both optical bias levels and input data pulses (Fig. 5). In principle, indefinitely extensible optical logic has been demonstrated. A generalized optical computer using these ideas may take the form shown in Fig. 6.

Conclusion

One may conclude that practical optical circuit elements have been demonstrated, that all binary logical operations can be performed and that the first digital optical circuits have been shown to be practical.

At this time the most favourable method of applying optical methods to digital optical computing seems to be the use of digital arrays of gates speaking (in an optical sense) to further arrays of gates with the possibility of fixed or variable intercom-

munications, fan-out and fan-in using free space propagation as efficient ways of 'optical wiring'. It remains to be shown that holding powers and switching energies are sufficiently small to allow operation with existing lasers and devices. Scaling from present experimental results, one may deduce that, for a visible wavelength device, a single logic element area of $\sim 1 \mu\text{m}^2$ should be practical, which, with a 250-ns switching time, should require 0.2 mW of holding power. Perhaps 10% of this would be absorbed. The 10 W available in the 514-nm argon ion line from a typical commercial laser would allow the operation of 5×10^4 gates assuming perfect optics. This would give the array a rate of information processing equivalent to 2×10^{11} logic operations per second assuming that all the gates are used simultaneously. A similar number of logic operations could be achieved by using small-gap semiconductors in the infrared where the electronic processes would allow faster switching but the packing density and device size would be greater. The optimization of the existing devices and the use of different materials may allow the synthesis of arrays capable of much higher data rates.

Even allowing for imperfections, which would reduce these numbers, the data rates are sufficiently high to encourage the exploration of their use for tasks where conventional one-dimensional sequential digital electronics finds difficulties, such as pattern recognition, artificial intelligence, sorting and specialized computational problems. Equally interesting will be applications such as power limiters, optical noise reduction and laser projectors and displays.

International development programmes

Several collaborating groups are actively involved in research in these areas. The principal effort in Europe is the European Joint Optical Bistability Program, a multinational project on optical logic circuitry and the basic physics of optical bistability involving eight universities and institutes in Britain, Belgium, West Germany, Italy and France. There are a further eight European laboratories associated with the programme. The research support scheme was established in 1984 by the Commission of the European Communities, through its Committee for the European Development of Science and Technology.

In the United States there are two main collaborative groups: The Optical Circuitry Cooperative is currently being established by the University of Arizona, which is involved with over 10 American companies with an interest in optical circuitry. The five-year programme is funded by these companies, the university and a NSF grant. The Pentagon recently announced the formation of a consortium of nine research organizations, including seven universities which will be funded as part of the Strategic Defense Initiative to produce an optical computer.

Japan's Optical Computer Group was formed in 1983 with the purpose of providing the opportunity for its 75 members to exchange information and ideas in the field of optical computing. The group has four to six meetings every year and publishes a newsletter (*Opcom news*) between meetings. The group is associated with the Optics Division of the Japanese Society of Applied Physics with members from Japanese universities and companies.

I thank Dr F. A. P. Tooley for help in preparing this manuscript as well as all members of the Heriot-Watt optical circuits research group and the participating laboratories of the European Joint Optical Bistability project of the Commission of the European Communities.

1. Franken, P. A., Hill, A. E., Peters, C. W. & Weinreich, G. *Phys. Rev. Lett.* **7**, 118-120 (1961).
2. Bloembergen, N. *Nonlinear Optics* (Benjamin, Menlo Park, 1965).
3. Mead, C. & Conway, L. *An Introduction to VLSI Systems* (Addison-Wesley, London, 1980).
4. *Opt. Engng* **24**, no. 1 J/F (1985).
5. Wherrett, B. S. & Smith, S. D. (eds) *Optical Bistability, Dynamical Nonlinearity and Photonic Logic* (Royal Society, London, 1984) and *Phil. Trans. R. Soc. A* **313**, 191-451 (1984).
6. *Meet. int. Comm. Opt.*, Sapporo (August 1984).
7. *Conf. opt. Soc. Am.*, Lake Tahoe (March 1984).
8. Szöke, A., Daneu, V., Goldhar, J. & Kurmit, N. A. *Appl. Phys. Lett.* **15**, 376-379 (1969).
9. Gibbs, H. M., McCall, S. L. & Venkatesan, T. N. C. *Phys. Rev. Lett.* **36**, 1135-1138 (1976).

10. Miller, D. A. B., Mozolowski, M. H., Miller, A. & Smith, S. D. *Opt. Commun.* **27**, 133-136 (1978).
11. Miller, D. A. B., Smith, S. D. & Johnston, A. *Appl. Phys. Lett.* **35**, 658-660 (1979).
12. Miller, D. A. B. & Smith, S. D. *Opt. Commun.* **31**, 101-104 (1979).
13. Gibbs, H. M. *et al. Appl. Phys. Lett.* **35**, 451-453 (1979).
14. Miller, D. A. B., Seaton, C. T., Prise, M. E. & Smith, S. D. *Phys. Rev. Lett.* **47**, 197-200 (1981).
15. Miller, D. A. B. *IEEE J. quant. Electron.* **QE-17**, 306-311 (1981).
16. Smith, S. D. *et al. Opt. Commun.* **51**, 356-362 (1984).
17. Smith, S. D. *et al. Opt. Engng* **24**(4) (July/August 1984).
18. Wherrett, B. S. *Appl. Opt.* (in the press).

Laser biology and medicine

V. S. Letokhov

Institute of Spectroscopy and Laser Technology Centre of the USSR Academy of Sciences, Moscow, Podolskii Rayon, Troitzk 142092, USSR

The substantial repertoire of laser radiation—its coherence, range of intensity and frequency, controllability of focal area and of beam length—has been comprehensively explored in the contexts of micro- and macro-diagnostics of cells, biochemical kinetics, therapy and surgery.

TWENTY-FIVE years have passed since the first laser, whose underlying principles were described by Basov, Prokhorov, Townes and Schawlow^{1,2}, was created³. The main applications of lasers were immediately forecast, but the widespread subsequent developments, such as diagnostic, therapeutic and surgical tools in biology and medicine, were not foreseen.

Biomedical laser applications have already been examined in a series of reviews⁴⁻⁸. The present review, timed to coincide with the 25th anniversary of the laser, surveys some recent developments in this field and outlines the new possibilities of laser techniques in disciplines such as photobiology and photo-medicine, concentrating on the main trends in development.

All biomedical laser applications are based on the interaction of laser light with biological substances. Such interaction causes a broad spectrum of effects, so lasers can be used in a wide range of fundamentally different applications (Fig. 1). In the simplest case, low-intensity laser light is absorbed, reflected or reradiated (as fluorescence) by the substance so that no changes occur within it. Such interactions form the basis for diagnostic laser applications in biology and medicine, both for diagnosis at the atomic or molecular level (spectral diagnostics) and at the macroscopic level (such as motion and shape of cells and organs).

Ultraviolet and visible radiation can excite electronic states in many biologically important molecules leading to photochemical transformations at the molecular level. Such processes occurring under the effect of light from incoherent sources are studied by photobiologists and find practical application in photomedicine^{9,10}. The use of laser light in these disciplines gives several benefits, first, it enhances the speed at which molecules are excited, making it possible to probe photobiochemical reactions on a picosecond timescale¹¹ and second, it excites higher electronic states in the target molecules through multiphoton processes.

High-intensity laser radiation can cause damage to matter through various mechanisms, for example, liquefaction, intensive heating, evaporation and dielectric breakdown. Such processes, rarely observed under the effect of incoherent light sources given their relatively low intensity, now underlie laser material processing. The same physical processes have proved effective in laser-beam surgery.

Fundamental to these diverse applications of laser light is its unique properties: (1) spatial coherence, allowing laser radiation to be focused and introduced into an optical fibre; (2) monochromaticity or temporal coherence, allowing laser spectral analysis and the resonance excitation of certain molecules; (3) high intensity (or temperature), providing a high concentration of light energy in the substance being irradiated, thus giving

rise to multiphoton processes, heating in the substance and other processes; (4) 'compressibility' into ultra-short pulses, allowing the study of very fast (subpicosecond) primary photoprocesses and the excitation of high-energy levels in molecules at a rate exceeding their relaxation rate¹²; (5) tunability in the range from the infrared to the ultraviolet, which permits selective excitation of certain molecular species. Of course a single laser can provide a combination of several of the above properties, although only one or two are essential to each particular application. I consider here the three areas of biomedical laser applications.

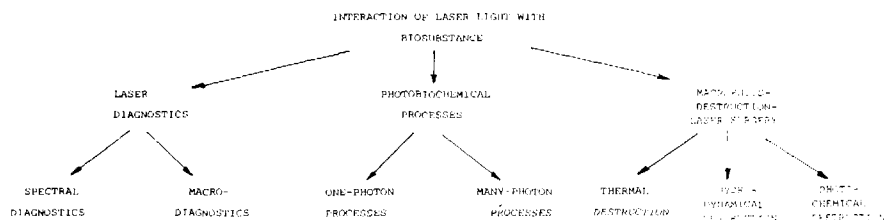
Laser diagnostics

Of importance in biology and medicine are microdiagnostics at the atomic and molecular level and macrodiagnostics at the level of cells and organs. Laser techniques are of great value in both cases.

Atomic and molecular diagnostics. Spectral analyses, such as trace-element analysis in studies of metabolism and toxicology, absorption and fluorescence techniques for the analysis of biological molecules, were performed extensively before the laser was devised. However, spectral analysis without the laser enables the detection of at best $\sim 10^{10}$ atoms or molecules of a particular species in a typical sample. Using lasers, ultra-sensitive analysis techniques^{13,14} were developed which make it possible to detect even single atom or molecule¹⁵. Furthermore the high spectral resolution and selectivity of laser spectroscopy techniques frequently allow the atomic content of samples to be determined directly, without resorting to preliminary separation into components and subsequent chemical analysis. (One example is direct resonance ionization detection of traces of aluminium in human blood¹⁶.) This highly sensitive and selective technique is universal in scope and provides a means of classifying the content of elements in various human organs (down to specific cells) and of revealing their role in human metabolism under both normal and pathological conditions.

This task is inseparable from the analysis of the environment (for example, flora and fauna, foodstuffs) for toxic and pathogenic materials, and the determination of pathways by which they enter the human organism. Here, laser-spectroscopy techniques are especially effective^{13,14}. Even the comparatively simple fluorescence analysis method is highly sensitive when used with the laser, in combination with chromatography. For example, the detection sensitivity for aflatoxins¹⁷ was about 10^{-10} g (10^{-12} mol) until the laser technique was developed; laser resonance photoionization of molecules and mass spectrometry increase the detection sensitivity by several orders of magnitude. Thus, for example, a detection sensitivity as high as 10^{-14} g has been reached for tryptophan in water¹⁸.

Fig. 1 Main trends of laser applications in biology and medicine, based on various degrees of laser light interaction with molecules, from diagnostics to destruction.



The absorption and fluorescence spectra of human tissues under normal conditions differ from those under pathological conditions, a phenomenon long used in diagnostics. The use of laser light, which can be easily delivered into the human body and focused on the desired spot, has extended the capabilities of such techniques. For instance, laser-fluorescence bronchoscopy is effective in early lung-cancer localization¹⁹. The technique uses a sensitizer, haematoporphyrin derivative (HpD), which concentrates predominantly in tumour cells. This compound absorbs light in the 400-nm region and fluoresces in the red region, allowing the detection of cancer at an early stage. Recently it has been shown that the fluorescence spectra of normal and atherosclerotic blood-vessel walls differ appreciably²⁰, a phenomenon which may prove valuable in operative diagnosis in blood-vessel surgery, an area of particular importance in view of the prospects for the development of laser blood-vessel surgery (laser angioplasty), discussed below.

Optoacoustic spectroscopy may, in future, be used to localize intraocular malignant melanomas²¹. The technique involves the generation of ultrasound at a spot inside the eye by focused laser radiation and detecting inhomogeneities by monitoring both the generation and propagation of the sound.

The development of fibre optics and the subsequent creation of fibre-optic sensors for biomedical applications²² have extended the capabilities of the laser spectral diagnosis of internal diseases. Such sensors have already been successfully used for fluorometric analysis, but possibilities here extend much further, because optical fibres allow transmission of light to the desired spot inside the human body not only in therapeutic but also in higher, surgical, radiation doses.

Laser radiation can be focused into a small spot of the order of the radiation wavelength, λ , in size. This property is being exploited in a range of techniques for the microspectral analysis of biological objects. For example, focusing an intense laser pulse onto the surface of a biological object can help to evaporate a $\sim 1\text{-}\mu\text{m}^3$ microsample and produce atomic and molecular ion species in the process; mass analysis of these species can provide information on the atomic and molecular composition of the sample²³. This principle underlies the commercial laser micromass analyser (LAMMA) manufactured by Leybold Heraeus.

Using a tunable laser radiation of a lower intensity allows the study of Raman scattering, absorption and fluorescence spectra of small areas of material without its being destroyed. This forms the basis, for example, of the laser microfluorometry of the processes occurring in single cells or organelles in the living state, the spatial resolution is as high as $0.3\text{ }\mu\text{m}$ and temporal resolution, 0.2 ns (ref. 24). Such a technique could be useful in the fluorescence mapping of genes. One possible technique for the direct observation of the primary structure of DNA is by means of a combination of the selective laser ionization of molecular chromophores with ion-field microscopy²⁵. This technique, which I and my colleagues have been working on for 10 years, would enable analysis of genetic information of individuals and early diagnosis of hereditary diseases.

Currently at the development stage are laser techniques for detecting rare radioactive isotopes, based on the isotope-selective detection of single atoms²⁶ and independent of the radioactive decay process itself. Such methods can be used as the basis for isotope analysis *in vivo* of metabolic pathways in an organism using rare isotope tracers of ultra-low radioactivity. In this context, it would be possible to conduct experiments, even at the cellular level, at a harmless radioactive dose.

Lasers can produce powerful ultra-short light pulses with a duration in the range 10^{-9} – 10^{13} s and, recently, in the femto-second range (down to $20\text{ fs} = 2 \times 10^{-14}$ s) (ref. 27). Picosecond light pulses have found wide applications in the investigation of primary processes involved in photosynthesis, vision and biochemical reactions involving haemoglobin, DNA and other biologically important molecules, which occur after absorption of a photon^{11,28}. This application of lasers has developed into a vast area of research, providing biochemists with direct and

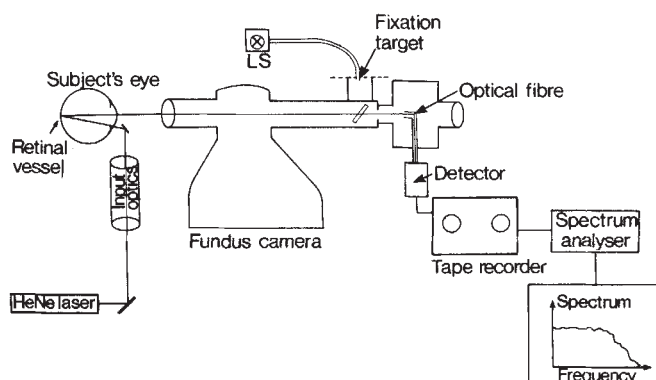


Fig. 2 Scheme of optical and electronic arrangement used to measure Doppler-shifted spectra from human retinal vessels. The basic optical component is a standard fundus camera modified with laser-input optics, internal fixation target and an optical fibre mounted in a scanning ocular (from ref. 32).

reliable information on ultra-fast primary photoprocesses for which only indirect data have been previously available.

Macrodiagnostics. The high monochromaticity and coherence of laser radiation allow measurement of the position, velocity, small displacements and shapes of various components of biological objects, that is, to perform macrodiagnostics for biomedical purposes.

One of the first effective biomedical laser applications was in flow cytometry, where the laser was used to effect the rapid analysis and separation of individual mammalian cells^{29,30} by accurately measuring their optical properties (particularly laser-induced fluorescence characteristics). The cells in suspension exist in a fast-moving, very thin laminar stream which is irradiated by a laser. When a cell of the desired species crosses the laser beam, it begins to fluoresce. The fluorescence triggers a system that diverts the droplet containing the cell. Such laser fluorocytometers are now available commercially; the next step will be to develop a technique for separating male and female cells by the slight differences in their fluorescence spectra resulting from differences between their sets of chromosomes.

Another effective laser application is the measurement of slow fluid flow velocities, particularly blood-flow velocity in vessels. The technique is based on the Doppler shift of the output frequency of a monochromatic laser (usually a He-Ne laser operating at $\lambda = 6,328\text{ }\text{\AA}$), which occurs as the laser radiation is scattered back from micron-sized particles moving in the blood flow. For example, at a flow of $v = 1\text{ cm s}^{-1}$ the frequency shift, $\Delta\nu$ of the He-Ne laser is $\Delta\nu = V/\lambda = 1.58\text{ kHz}$ and can be detected by monitoring the beat signal on a photodetector receiving both the incident and reflected (back-scattered) radiation. Laser Doppler velocimeters allow for continuous non-invasive blood-flow velocity measurements at the capillary level³¹. As an example, Fig. 2 illustrates schematically the direct measurement of retinal blood-flow velocity in the eye³². The eye is especially convenient for such measurements as its natural optics make it easy to deliver light to the desired microcapillary with an irradiation region diameter of only $100\text{ }\mu\text{m}$. However, the development of fibre optics has made this technique applicable to other blood vessels³³. These fibres enable the measurement of blood-flow velocity profiles with a resolution of 0.1 mm or better, and velocity variations lasting $\sim 0.01\text{ s}$ or less²². By such means it will be possible to measure microcirculatory blood flows with high resolution using an optical fibre brought into contact with the tissue of interest. The development of compact diode lasers should eventually make this technique available for general medical applications. Moreover, the capabilities of the laser-Doppler technique extend much further, for example, it can be used to measure the mobility of bacteria and spermatozoa and to perform electrophoresis of cells and organelles³⁴.

With the advent of lasers, one of the most important pre-laser discoveries in optics, holography³⁵, has become a powerful

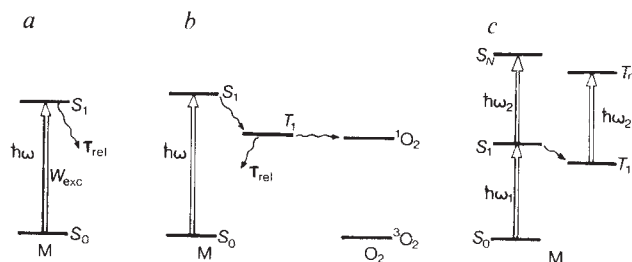


Fig. 3 Methods of photochemical modifications of biomolecules. *a*, One-photon excitation and subsequent photochemical reaction (phototherapy of neonatal jaundice by means of photodecomposition of bilirubin). *b*, One-photon excitation of sensitizer molecule, *M*, and transfer of excitation to modifiable molecule or molecules participating in modification photodynamical effect (tumour phototherapy; the oxygen-acceptor of excitation participates in oxidation). *c*, Two-photon excitation of highly lying triplet, T_1 , or singlet, S_1 , states and following photomodification or transfer of excitation to other molecules (nucleic acids in water).

diagnostic method. Using holography it is possible to obtain three-dimensional images of biological objects; their contours can be mapped (holographic topography) and their deformations can be analysed on a real timescale (holographic interferometry). These new possibilities have been influential in many areas of medicine, including orthopaedics, radiology, ophthalmology, urology and otology³⁶.

Femtosecond light pulses are compressed in space (a pulse duration of 50 fs = 5×10^{-14} s corresponds to a pulse length of 15 μ m in air). Thus, the possibility arises of direct optical measurement of small distances on the principle of a pulsed optical radar. This technique is applicable to investigations with micrometre-scale spatial resolution into biological systems such as the interior of the eye or skin³⁷. When combined with fibre optics, the technique may have interesting applications in studying human internal organs.

Molecular photobiology and photomedicine

Within this vast area lie those laser applications which make use of photochemical processes resulting from photo-excitation of biological molecules (Fig. 1). To outline what laser light can provide in this area, I consider general schemes of molecular photoexcitation following adsorption of one or more photons (Fig. 3). When being acted on by low-intensity light, a molecule can absorb no more than one photon because the excitation rate $W_{\text{exc}} = \sigma I$ is much lower than the relaxation rate $1/\tau_{\text{rel}}$ of the excited molecule (σ is the excitation cross-section, I the radiation intensity in photons/cm²s, and τ_{rel} the excitation-relaxation time). The excited molecule either takes part in a chemical reaction (Fig. 3*a*) or transfers its excitation to another molecule participating in a chemical transformation (Fig. 3*b*). However, under the effect of short laser pulses of intensity I is so high that $W_{\text{exc}} \approx 1/\tau_{\text{rel}}$, the molecule can absorb a second photon either from a triplet state or from an excited singlet state (Fig. 3*c*). The excited molecule again can suffer both a photochemical transformation and transfer excitation to another photochemically active molecule. The first two schemes form the basis for linear, one-photon photobiology⁹, while the third (and, perhaps, schemes involving the participation of more than two photons) underlies nonlinear, multiphoton photobiology^{12,38}. These two approaches impose different requirements on the laser intensity.

One-photon excitation. One-photon photobiochemical processes, such as phototherapy of neonatal jaundice (hyperbilirubinaemia)³⁹ (Fig. 3*a*) phototherapy and photochemotherapy of various skin diseases (Fig. 3*a, b*)⁴⁰ and photochemotherapy of cancer with the aid of the haematoporphyrin derivative HpD⁴¹ are well known in photobiology⁹.

The last case is a classical example of phototherapy following the scheme of Fig. 3*b*. The HpD molecule tends to accumulate

predominantly in malignant tissues. At the same time, HpD is readily activated with visible light and, once excited, transfers its excitation through a triplet state to the oxygen molecules present in the tissues. The oxygen molecules excited to a singlet state ($^1\text{O}_2$) are chemically active and cytotoxic. The entire mechanism is termed the photodynamic effect⁴².

The light intensities required for all photoprocesses are not very high ($<1 \text{ W cm}^{-2}$), and so they can be performed under the effect of light from incoherent sources. Nevertheless, the use of laser light here is extremely useful for the following reasons. First, laser radiation is easily delivered to the internal organs through optical fibres without significant loss of intensity. This considerably extends the capabilities of phototherapy, allowing one, in principle, to change from the treatment of external organs to that of any internal organ⁴.

Second, the high intensity of laser radiation ensures a high rate of photochemical processes, irrespective of the excitation of high-lying electronic states. This laser light provides a therapeutic dose over a briefer time span than is possible with radiation from conventional sources. It is for this reason that lasers have given a fresh impetus to the development and improvement of established cancer photochemotherapy techniques using ultraviolet and visible radiation⁴¹.

Third, the necessary irradiation dose can be provided by a single short pulse, which is of particular use in those cases where the effect of light on subsequent secondary chemical reactions must be excluded, for example, experiments on ultraviolet photoreactions of psoralen with DNA, which proceed according to the scheme of Fig. 3*b*⁴³.

Psoralen can intercalate between the pairs of bases of double-stranded nucleic acids, forming covalent adducts with the pyrimidine bases on the absorption of a long-wave ultraviolet quantum. If there is another pyrimidine base on the opposite side, it can also form a second covalent adduct on absorption of a second photon by psoralen. The diadduct thus formed acts as an interstrand crosslink. Where irradiation lasts long enough there is always a large proportion of lethal cross-links which make it difficult to conduct experiments *in vivo*. Irradiation with a single short ultraviolet pulse, however, can be made to yield only monoadducts, as the psoralen molecule can absorb a second photon only after relaxing to the ground state through a triplet state; this takes longer than the pulse duration.

It is precisely by this method that the proportion of mono- and diadducts in the psoralen-DNA photochemical reaction have been controlled following irradiation with a short (15 ns) pulse of an N_2 -laser ($\lambda = 337 \text{ nm}$)⁴⁴. When using a pulsed radiation of a shorter wavelength (a KrF laser, $\lambda = 248 \text{ nm}$), which could be absorbed by the nucleic acid bases themselves, the formation of crosslinks occurred between RNA polymerase and DNA⁴⁵. Again, the use of short ultraviolet laser pulses is advantageous. Photodamage can be caused rapidly, but no structural changes are possible during the time of a single pulse.

Recently, a low-intensity (non-coagulative) He-Ne laser radiation ($\lambda = 632.8 \text{ nm}$) has been successfully used in clinical practice for treating wounds and trophic ulcers^{46,47}. The therapeutic effect of red laser light is frequently explained by its coherence. This argument seems unconvincing, however, as the molecular excitation rate W_{exc} is at least a factor of 10^{10} lower than the rate at which the excited molecules lose their coherence in a condensed medium at normal temperature⁴⁸. Experiments are under way to study this biostimulative effect at the level of mammalian cells, yeasts and bacteria, which have shown that the effect is present regardless of the coherence of the light used, provided its wavelength lies in the appropriate range, 600–850 nm⁴⁹. Thus, this biostimulative effect has no specific relation to the laser and should be considered simply a photobiological effect with useful applications to phototherapy.

Multiple-quantum excitation. The idea that multiple-quantum excitation of biological molecules can be produced by ultra-short laser pulses of low energy but high-peak power was formulated⁵⁰ immediately after such processes had been observed in the gaseous phase (see refs 12, 38). An intrinsic difference in the

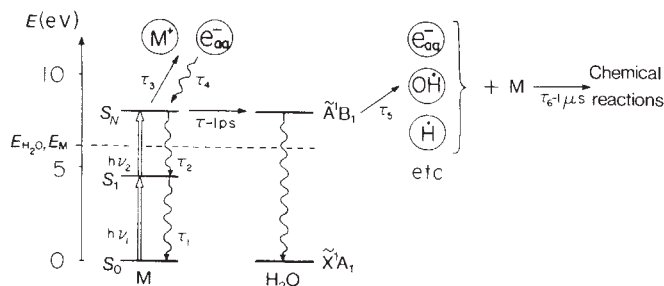


Fig. 4 Simplified diagram of primary photoprocesses occurring in picosecond two-quantum excitation of biological molecules (thymine, a DNA base) in water. Absorption of two ultraviolet photons results in excitation of the S_N Rydberg state lying above the ionization threshold of both the dissolved molecule and water. $\tau_1 \approx 10^{-12}$ and $\tau_2 \approx 10^{-13}$ s are the non-radiative relaxation times for the states S_1 and S_N ; $\tau_3 \approx 10^{-11}$ and $\tau_4 \approx 5 \times 10^{-12}$ s are the times of formation and recombination of an electron-ion pair; $\tau = 10^{-12}$ s is the time of electronic energy transfer from the dissolved molecule to water. Subsequent ionization and dissociation of water ($\tau_5 \approx 10^{-13}$ – 10^{-12} s) leads to the formation of reactive radicals. Chemical reactions involving water radicals ($\tau_6 \approx 10^{-6}$ s) are the main photomodification channel for the dissolved molecules.

case of biological molecules in solution is the rapid non-radiative decay of their excited states. The situation is less critical if the molecule relaxes through a triplet state with a comparatively long lifetime (10^{-6} – 10^{-9} s). In this case, a two-step excitation can be effected by means of nanosecond laser pulses (Fig. 3c). However, for the molecule to absorb a second photon from a singlet state, its excitation rate must be close to the relaxation rate that lies in the picosecond range $1/\tau_{rel} \approx 10^{11}$ – 10^{12} s⁻¹. Such a two-step excitation can, therefore, be effected only by means of ultra-short laser pulses.

The first successful experiments on the two-quantum excitation of such molecules with picosecond laser pulses and the ensuing photoreactions were conducted with aqueous solutions of nucleic acid bases^{51,52}. As a result of such irradiation, the molecule absorbs two ultraviolet photons ($\lambda = 266$ nm) with a total energy of $2h\nu = 9$ eV which exceeds the ionization and dissociation limit of the molecules of the solvent, water. (Note that the direct one-quantum excitation to the same state ($\lambda = 133$ nm) is impossible because of the strong absorption of water in the ultraviolet.) The highly excited molecules may themselves undergo ionization with a yield that largely depends on the competition between the two-quantum photoionization and recombination processes (Fig. 4). Another possibility is the transfer of excitation to the molecules of the solvent, H₂O, which are subsequently ionized and dissociate⁵³. As a result of this process, the water molecules are decomposed into radicals that react with the dissolved biological molecules. Therefore, when using the two-quantum method to excite high-lying electronic states of biomolecules (for example, thymine) one can observe photobiochemical reactions that usually occur in solution in response to ionizing radiation⁵⁴. However, the ultraviolet picosecond photolysis of biomolecules has several advantages over their γ -radiolysis.

One principal distinction between two-quantum photolysis and γ -radiolysis is that the solvent molecules in the former case are dissociated only near the chromophore molecules which absorb laser radiation, and not within the entire irradiated volume, as is the case with γ -radiolysis. This fact is used to advantage in the laser therapy of tumours, for it provides the possibility of locally producing water radicals that kill tumour cells but have no effect on nearby normal cells, by means of specific sensitizer molecules.

Two-quantum photobiochemical reactions have been observed at various levels of organization in biological systems (for example, in viruses, yeasts and cells)^{55,56} (see ref. 57 for a review). This approach will apply in genetic engineering, especially in combination with microirradiation of cells⁵⁶. For

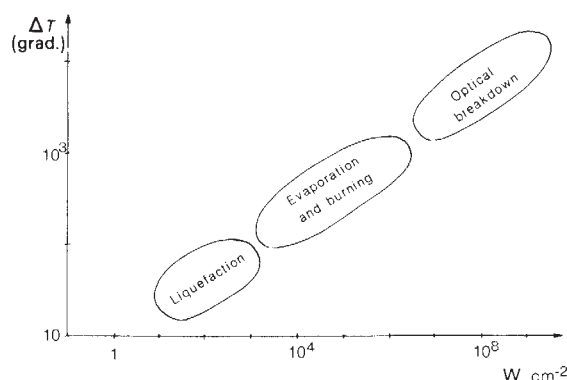


Fig. 5 Various degrees of heating of a substance with the laser radiation that underlie the various laser surgery techniques in ophthalmology.

instance, various types of molecular damage (for example, breaks, crosslinks) have been observed in nucleic acids *in vivo*^{55,57}. This allows the investigation of the spatial organization and functions of complex molecules such as proteins and nucleic acids, especially where effective methods for identifying damage sites are available⁵⁸.

In addition to the direct two- (multiple-) quantum process in biological molecules (following the scheme in Fig. 3c), one can use laser action with the help of sensitizer molecules, which absorb laser photons more effectively. These sensitizers are generally dyes which can form complexes with biological molecules. Once the sensitizer molecule has absorbed enough energy, the biological molecule suffers damage at the spot where the complex has been formed. This method has been investigated in experiments with nucleic acid molecules^{59,60}. For example, polyadenylic acid molecules undergo scission at the point of complexing with the dansyl derivative of oligothymidilate (DDO) under the effect of a powerful N₂-laser pulse. This phenomenon is caused by the two-quantum excitation of DDO and subsequent transfer of excitation to the nucleic acid⁶⁰. It has been suggested that molecular sensitizers might be used to disrupt nucleic acid molecules at definite, pre-selected points^{60,61}.

The two-quantum excitation of high-lying electronic states is expected to increase considerably the yield of photoreactions compared with the known one-quantum reactions proceeding through triplet states, because the singlet-triplet conversion yield is in many cases much less than 100%. Such an increase in the yield of a photochemical reaction has been observed in experiments on the excitation of porphyrin molecules in solution⁶². Changing from a low-intensity excitation (<1 W cm⁻²) to a high-intensity excitation (10^9 W cm⁻²) increased the quantum yield of photomodification of porphyrin by a factor of 100. Similar experiments with a haematoporphyrin derivative, conducted with the N₂-laser, showed⁶³ that the primary mechanism responsible for the destruction of this molecule was its two-step excitation. Such excitation produces a stronger cytotoxic effect than low-intensity irradiation.

Summarizing the first results of multi-quantum excitation of biological molecules, we are close to the development of new effective nonlinear laser phototherapy techniques using ultra-short laser pulses capable of producing considerable photochemical effects at an average power too low for thermal effects. In particular, a wide-ranging phototherapy method will probably be developed, based on the transfer of excitation to water molecules⁵⁴, which will extend the capabilities of the existing radiotherapy techniques.

Laser surgery

The use of lasers in surgery is based mainly on the absorption of a substantial amount of light energy in the area being operated on, which causes the heating of the irradiated region and its subsequent destruction by one of several possible mechanisms,

depending on the irradiation conditions. Such photodestruction of the tissue requires no contact, allowing absolute sterility for laser surgery. Second, by properly selecting the laser wavelength, one can attain a certain selectivity of photodestruction, provided that the tissue to be removed differs in absorption spectrum from the surrounding tissues. Third, the thermal effects can be varied over a very wide spectrum, from liquefaction without evaporation⁶⁴ under slight heating to destruction by hydrodynamic shock wave resulting from local intense pulsed heating⁶⁵. Finally, the laser irradiation region can be localized down to a few fractions of a micrometer, which enables organs to be operated on and microsurgery at the subcellular level⁶⁶.

The vast field of laser surgery is summarized briefly here, using ophthalmology as an example (Fig. 5). A continuous wave CO₂-laser radiation can be used to heat the surface layer of the vitreous humour of the eye from the ambient temperature to the gel liquefaction temperature and then to aspirate the liquefied area. In this case, the energy consumed in liquefaction (53 cal g⁻¹) is about one-tenth of that required for evaporation (576 cal g⁻¹). By increasing the radiated power, intense heating and coagulation of the irradiated region can be achieved. This is used, for example, in the treatment of proliferative diabetic retinopathy by photocoagulation⁶⁷. Focused laser radiation can be used to heat intensely and evaporate small areas inside the eye to open up a pathway into Schlemm's canal in treating glaucoma⁶⁸. In such microsurgical operations, it is essential that photodamage is strictly localized, to avoid intense heating and damage to the surrounding tissues. In this case it is advantageous to use a short, powerful laser pulse which, because of the nonlinear effect occurring at the focal point, can be absorbed even in a transparent tissue. This causes intense local heating and destruction in the area affected by the shock wave resulting from high hydrodynamic pressures; the surrounding tissue suffers no appreciable heating⁶⁵.

The field of application for laser-surgery techniques has been extended with the development of methods to deliver a surgical radiation dose (from a few watts to a few tens of watts) through an optical fibre passed through blood vessels to internal areas of the human body. Laser-angioplasty techniques^{69,70}, in which a fibre optics catheter is used to deliver a laser radiation to atherosclerotic obstructions (plaques) are particularly promising. To evaporate a plaque, it is sufficient to irradiate it by an argon laser of 3–4 W in power output for a few seconds, and the evaporation of a calciferous plaque requires 30–40 s of irradiation with the same laser. The destruction of an obstruction using infrared or visible CW radiation occurs according to a thermal mechanism, and to prevent thermal damage to the vessel walls, one is restricted to piercing only a small-diameter passage. Of interest, therefore, is the change from thermal destruction with CW laser radiation to photochemical ablation of plaques by pulsed ultraviolet radiation. In such photochemical actions, the radiation energy is expended directly in the breakdown of chemical bonds, followed by the destruction of the irradiated material as a result of volatile products⁷¹. Some preliminary reports on the effect of ultraviolet-laser radiation on atherosclerotic plaques *in vitro* show that the mechanism of action here is distinct from the thermal mechanism in visible and infrared irradiation⁷². This conclusion is in line with the results of studying the mechanism and products of plaques destruction by ArF ($\lambda = 193$ nm)⁷³, XeCl ($\lambda = 308$ nm) excimer and other lasers⁷⁴. The specific photo-ablation energy amounts to around 250 cal g⁻¹, that is, it is less than the specific thermodestruction energy⁷⁴. Thus, another promising method can be added to the list of potential photodestruction techniques in laser surgery.

Finally, sharp focusing of laser radiation (microbeam techniques) makes it possible to alter selectively some part of a subcellular organelle in a single living cell⁷⁵. The region of such alterations can be only a quarter of a micrometre in length. The laser wavelength and pulse duration can be varied over a broad range, allowing for photodestruction conditions, from thermal to multiquantum photochemical. This opens up new possibilities

in cell and developmental biology, enabling chromosome micro-surgery, surgery of mitotic organelles and cytoplasm⁷⁵, for example. The capabilities of the technique extend into genetic engineering, as it provides experimentalists with a laser scalpel enabling a submicron hole to be pierced in the wall of a cell. The desired gene from the surrounding solution can enter the cell through this hole which is then healed again in a few fractions of a second⁷⁶ (J. Wolfrum, personal communication).

Thus, laser biology and medicine now use all the characteristics of laser light as well as all types of its interaction with biological structures. The many fascinating and rapidly developing laser applications are sure to become more widespread and diverse in the future.

I thank Drs T. Karu and D. Nikogosyan for their assistance and fruitful discussions in the course of preparation of the manuscript.

- Basov, N. G., Prokhorov, A. M. & Townes, C. H. *Nobel Lectures* (Nobel Foundation, Stockholm, 1964).
- Schawlow, A. L. & Townes, C. H. *Phys. Rev.* **112**, 1940–1949 (1958).
- Maiman, T. H. *Nature* **187**, 493 (1960).
- Fine, S. & Klein, E. *Laser Focus* **5**, 28–40 (1969).
- Wolbarsht, M. L. (ed.) *Laser Applications in Medicine and Biology* Vols 1–3 (Plenum, New York, 1971, 1974, 1977).
- Berns, M. W. in *Photochemical and Photo-Biological Reviews* Vol. 2 (ed. Smith, K. C.) 1–38 (Plenum, New York, 1977).
- Pratesi, R. & Sacchi, C. A., eds *Lasers in Photomedicine and Photobiology* (Springer, Berlin, 1980).
- Goldman, L. (ed.) *The Biomedical Laser. Technology and Clinical Applications* (Springer, Berlin, 1981).
- Smith, K. C. (ed.) *The Science of Photobiology* (Plenum, New York, 1977).
- Giese, A. C. (ed.) *Photophysiology* Vol. 1–8 (Academic, New York, 1964–73).
- Rentzepis, P. M. *Science* **202**, 174–182 (1978).
- Letokhov, V. S. *Nature* **305**, 103–108 (1983).
- Kliger, D. S. (ed.) *Ultrasensitive Laser Spectroscopy* (Academic, New York, 1983).
- Letokhov, V. S. (ed.) *Laser Analytical Spectrochemistry* (Hilger, London, in the press).
- Letokhov, V. S. *Optica Acta* (in the press).
- Bekov, G. I., Letokhov, V. S. & Radayev, V. N. *Laser Chem.* **5**, 11–17 (1984).
- Bradley, A. B. & Zare, R. N. *J. Am. Chem. Soc.* **98**, 620 (1976).
- Antonov, V. S., Letokhov, V. S. & Shibanov, A. N. *J. opt. Soc. Am. B* (in the press).
- Doiron, D. R. & Profio, A. E., in *Lasers in Photomedicine and Photobiology* (eds Pratesi, R. & Sacchi, C.) 92–95 (Springer, Berlin, 1980).
- Kittrell, C., Willett, R. L., de los Santos, C., Rat'liff, N. B. & Feld, M. S. *Appl. Opt.* (in the press).
- Landers, M. B. & Wolbarsht, M. L. in *The Biomedical Laser* (ed. Goldman, L.) 117 (Springer, Berlin, 1981).
- Peterson, J. I. & Vurek, G. G. *Science* **224**, 123 (1984).
- Kaufmann, R. & Hillenkamp, F. & Wechsung, R. *Med. Progr. Technol.* **6**, 106 (1979).
- Andreoni, A., Longini, A., Sacchi, C. A. & Svelto, O. in *The Biomedical Laser* (ed. Goldman, L.) 69–83 (Springer, Berlin, 1981).
- Letokhov, V. S. in *Tunable Lasers and Applications* Vol. 3 (eds Mooradian, A., Jaeger, T. & Stoketh, R.) 122–139 (Springer, Berlin, 1976); *Comm. At. molec. Phys.* **11**, 1–12 (1981).
- Balykin, V. I. et al. *Laser Spectroscopy*, VI (eds Wever, H. P. & Lüthy, W.) 103–105 (Springer, Berlin, 1983).
- Shank, C. V. *Science* **219**, 1027–1031 (1983).
- Alfano, R. R. (ed.) *Biological Events Probed by Ultrafast Laser Spectroscopy* (Academic, New York, 1982).
- Mullaney, P. F., Steinkamp, J. A., Crissman, H. A., Cram, L. S. & Holm, D. H. in *Laser Applications in Medicine and Biology* (ed. Wolbarsht, M. L.) Vol. 2, 151 (Plenum, New York, 1974).
- Salzman, G. C. in *The Biomedical Laser* (ed. Goldman, L.) 33–53 (Springer, Berlin, 1981).
- Stern, M. D. *Nature* **254**, 56–58 (1975).
- Riva, C. E., Feke, G. T. in *The Biomedical Laser* (ed. Goldman, L.) 135–161 (Springer, Berlin, 1981).
- Tanaka, T., Benedek, G. B. *Appl. Opt.* **14**, 189–196 (1975).
- Sattelle, D. B., Lee, W. I. & Ware, B. R. eds *Biomedical Applications of Laser Light Scattering* (Elsevier, Amsterdam, 1982).
- Gabor, D. *Nature* **161**, 777 (1948).
- von Bally, G. (ed.) *Holography in Medicine and Biology*, Vol. 18, (Springer, Berlin, 1979).
- Fujimoto, J. G. et al. *Adv. Prog. Conf. Lasers Electro-Optics* **45** (Baltimore, 1985).
- Letokhov, V. S. *Nonlinear Laser Chemistry, Multiple Photon Excitation* Vol. 22 (Springer, Berlin, 1983).
- Brown, A. K. & McDonagh, A. F. *Adv. Periatr* **27**, 341–389 (1980).
- Parrish, J. A. *J. invest. Dermat.* **77**, 167 (1981).
- Dougherty, T. J., Thoma, R. E., Boyle, D. G. & Weishaupt, K. P. in *Lasers in Photomedicine and Biology* (eds Pratesi, R. & Sacchi, C. A.) 67–75 (Springer, Berlin, 1980).
- Spikes, J. & MacKnight, F. *Ann. N.Y. Acad. Sci.* **171**, 149–162 (1970).
- Hearst, J. E. in *Chemical and Biochemical Applications of Lasers* (ed. Moore, C. B.) Vol. 4 389–403 (Academic, New York, 1979).
- Johnston, B. H., Johnson, M. A., Moore, C. B. & Hearst, J. E. *Science* **197**, 906–908 (1977).
- Harrison, C. A., Turner, D. H. & Hinkle, D. C. *Nucleic Acids Res.* **10**, 239–281 (1982).
- Gamaleya, N. F. in *The Biomedical Laser* (ed. Goldman, L.) 271 (Springer, Berlin, 1981).
- Mester, E. et al. in *Medizinische Erfahrungen in Forschung und Praxis* (ed. A. Kaul) 117 (De Gruyter, Amsterdam, 1976).
- Lobko, V. V., Karu, T. J. & Letokhov, V. S. *Biophysica (Russia)* **30**, 380–388 (1985).
- Karu, T. J., Kalendo, G. S., Letokhov, V. S. & Lobko, V. V. *Il nuovo Cimento D1*, 828–840 (1982); **D3**, 309–318 (1984); **D3**, 319–325 (1984).
- Letokhov, V. S. *J. Photochem.* **4**, 185–191 (1975).
- Kryukov, P. G., Letokhov, V. S., Matveev, Yu. A., Nikogosyan, D. N. & Sharkov, A. V. in *Picosecond Phenomena* Vol. 4 (eds Shank, C. V., Ippen, E. P. & Shapiro, S. L.) 158–166 (Springer, Berlin, 1978).
- Kryukov, P. G. et al. *Chem. Phys. Lett.* **61**, 375–379 (1979).
- Nikogosyan, D. N., Oravsky, A. A. & Letokhov, V. S. *Proc. USSR Acad. Sci.* (in the press).
- Oravsky, A. A., Nikogosyan, D. N. & Letokhov, V. S. *Lasers in the Life Sci.* (in the press).

55. Gurzadyan, G. G. *et al. Photochem. Photobiol.* **33**, 835-838 (1981).
56. Calmettes, P. P. & Berns, M. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7197-7199 (1983).
57. Nikogosyan, D. N. & Letokhov, V. S. *Revista nuovo Cimento* **6**, ser. 3, No. 8 (1983).
58. Brash, D. E. & Haseltine, W. A. *Nature* **298**, 189-192 (1982).
59. Andreoni, A., Cubeddu, R., de Silvestri, S., Laporta, P. & Svelto, O. *Phys. Rev. Lett.* **45**, 431 (1980).
60. Benimetskaya, L. Z. *et al. FEBS Lett.* **163**, 144-149 (1983).
61. Barashev, P. P., Demidov, V. V., Tal'rose, V. L. & Trofimov, V. I. *Proc. USSR Acad. Sci.* **266**, 1486-1490 (1982).
62. Karu, T. J., Kryukov, P. G., Ketokhov, V. S., Matveetz, Yu. A. & Semichishen, V. A. *Appl. Phys.* **24**, 245-247 (1981).
63. Andreoni, A., Cubeddu, R., de Silvestri, S., Laporta, P. & Svelto, O. *Chem. phys. Lett.* **88**, 37-39 (1982).
64. Bridges, T. J. *et al. Science* **219**, 1217-1219 (1983).
65. Mainster, M. A., Sliney, D. H., Belcher, C. D. III, & Buzney, S. M. *Ophthalmology* **90**, 973-991 (1983).
66. Berns, M. W., Olson, R. S. & Rounds, D. E. *Nature* **221**, 74-75 (1969).
67. Wolbarsht, M. L. & Landers, M. B. III. *Appl. Opt.* **18**, 1518-1526 (1979).
68. Krasnov, M. M. *Am. J. Ophth.* **75**, 674 (1975).
69. Lee, G., Ikeda, R. M., Kozina, J. & Mason, D. T. *Am. Heart J.* **102**, 1074-1077 (1981).
70. Choy, D. S., Stretzer, S. H., Rotterdam, H. Z. & Bruno, M. S. *Am. J. Card.* **50**, 1209 (1982).
71. Srinivasan, R. & Leigh, W. J. *J. Am. chem. Soc.* **104**, 6784 (1982).
72. Grundfest, W. *et al. Circulation* **70**, Suppl. II (1984).
73. Kinsker, R., Srinivasan, R., Wynne, J. J. & Alonso, D. R. *Lasers in Surgery and Medicine* (in the press).
74. Abil'sitov, G. A. *et al. Kvantovaya Elektronika* (in the press).
75. Berns, M. W. *et al. Science* **213**, 505-513 (1981).
76. Tsukakoshi, M. *et al. Appl. Phys. B* **35**, 135-140 (1984).

ARTICLE

Loss of heterozygosity in three embryonal tumours suggests a common pathogenetic mechanism

Alex Koufos^{*†}, Marc F. Hansen^{*}, Neal G. Copeland^{*†}, Nancy A. Jenkins^{*†}, Beatrice C. Lampkin[†] & Webster K. Cavenee^{*†}

^{*} Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267-0524, USA

[†] Division of Hematology/Oncology, Children's Hospital Medical Center, Cincinnati, Ohio 45267, USA

Children with the Beckwith-Wiedemann syndrome have a greatly increased potential for the specific development of the embryonal tumours hepatoblastoma, rhabdomyosarcoma and Wilms' tumour. Data obtained with molecular probes suggest that the association between these disparate, rare tumour types reflects a common pathogenetic mechanism that entails the somatic development of homozygosity for a mutant allele at a locus on human chromosome 11.

EMBRYONIC tumours in children may result from a developmental disturbance during organogenesis that causes arrest at an early stage in the normal differentiation process to adult tissue¹. This embryonal tissue continues to proliferate inappropriately at this early stage and produces a mass of immature tissue that has been simultaneously termed an embryonal tumour and an organismal malformation². Thus, these tumours can reasonably be considered to comprise one class of congenital malformations. The high-frequency association among some of these malformations has allowed clinical classifications of the developmental anomaly syndromes³. One example of this clinical grouping is the autosomal dominant Beckwith-Wiedemann syndrome (BWS)^{4,5}, which is composed of multiple congenital malformations coupled with a concurrent high risk for the development of specific rare childhood tumours. More than 10% of all individuals with BWS will develop one of the specific oncogenic malformations, which include Wilms' tumour, hepatoblastoma, rhabdomyosarcoma and adrenal carcinoma, in association with the growth excess disorders that characterize the syndrome^{5,6}. Even those individuals with the incomplete, partially penetrant form of the syndrome who display only a subset of these physical anomalies have increased risk for the development of specific malignancies⁶. In fact, these patients may develop tumours of more than one type simultaneously, such as the appearance of Wilms' tumour with adrenal carcinoma⁷ or adrenal adenoma⁸. Similar simultaneous development of more than one tumour type in patients with no readily apparent organismal anomalies has also been described. These include the development of Wilms' tumour with rhabdomyosarcoma⁹ or hepatoblastoma¹⁰, hepatoblastoma with adrenal adenoma¹¹, and rhabdomyosarcoma with adrenal carcinoma¹².

The suggestion of the ontogenetic relatedness of these tumour types is also supported by histopathological analyses. The embryonal tissue from which Wilms' tumour seems to arise

consists of the metanephric blastema that has failed to undergo the normal maturation process¹³. A common aetiology for Wilms' tumour, hepatoblastoma and rhabdomyosarcoma is suggested by the ability of these embryonal tumours to undergo divergent differentiation to form a tissue that does not occur in the normal end-stage organ. For example, fetal rhabdomyomatous nephroblastoma is a variant Wilms' tumour comprising striated muscle of the kidney¹⁴. The rhabdomyoblasts appear to arise by divergent or aberrant differentiation from the metanephric blastema¹⁴. Rhabdomyomatous tissue has also been observed in some hepatoblastomas¹⁵. Additionally, heterotropic primary tumours apparently consisting of adrenal carcinoma and rhabdomyosarcoma have been observed in surgically excised livers^{16,17}.

It has been reported that the expression of recessive mutant alleles at the *WAGR* locus on human chromosome 11 band p13 and the *Rb-1* locus on human chromosome 13 band q14 are involved in the development of Wilms' tumour¹⁸⁻²¹ and retinoblastoma²², respectively. Expression of the initial predisposing recessive mutation at these loci is effected by abnormal somatic chromosomal segregation events whose net result is the complete loss of the dominant balancing wild-type allele. The mitotic mechanisms that accomplish this removal included nondisjunctional loss and mitotic recombination. The results of these mitotic events are consistent with the production of a cell homozygous for the mutant allele at the retinoblastoma or Wilms' tumour locus. Evidence in support of this interpretation was recently obtained by the demonstration that the chromosome 13 homologue remaining in familial retinoblastoma tumours was always derived from the affected parent²³.

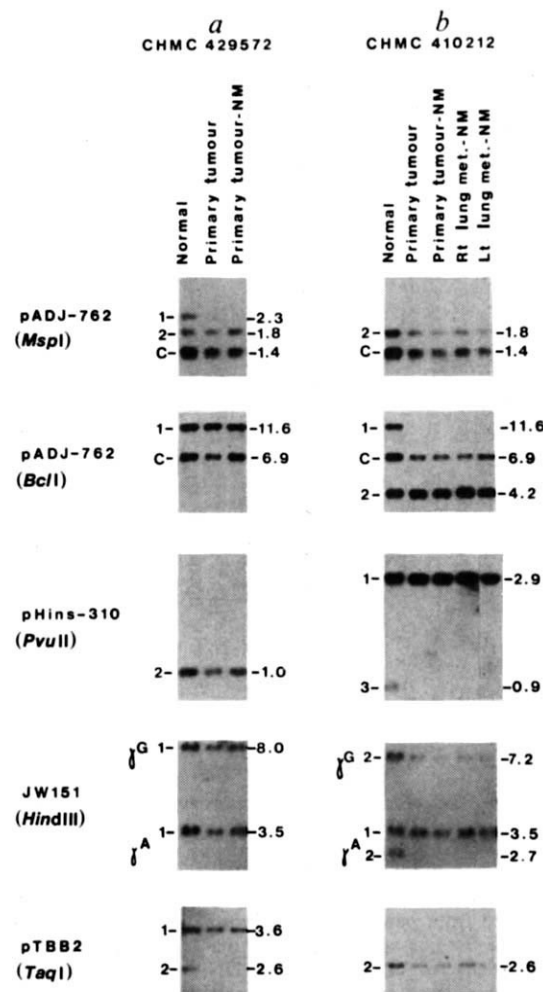
The clinical association of Wilms' tumour with other specific rare tumours in the presence of the growth excess syndromes, the development of more than one rare tumour in the same individual, and the presence of heterotropic tissue in these tumours could be simply circumstantial. Alternatively, these clinical associations may reflect a common aetiological event so that each of the developmental anomalies, including the tumours, arises after mutation of the same locus. Such mutations

[‡] Present address: Mammalian Genetics Laboratory, LBI Basic Research Program, NCI-Frederick, Maryland 21701, USA.

|| To whom correspondence should be addressed.

Fig. 1 Loss of constitutional heterozygosity at loci on human chromosome 11p in hepatoblastoma; *a*, patient A; *b*, patient B.

Methods. DNA was isolated from normal tissue (peripheral leukocytes), surgical specimens of primary tumour or from a xenograft of the primary tumour passaged twice in immunodeficient nude mice (NM), as described previously^{18,22,26}. DNA was also isolated from xenograft tumours derived from right and left lung metastases (passages 3 and 1, respectively) removed from patient B following thoracotomy and wedge resection to remove metastatic hepatoblastoma nodules. DNA from each of these tissues was digested to completion with the indicated restriction endonucleases, subjected to agarose gel electrophoresis and transferred to Zeta-por membranes (AMF-Cuno) in 1.5 M NaCl/0.15 M sodium citrate. The membranes were prehybridized at 42 °C for 16 h in 0.75 M NaCl/0.075 M sodium citrate, 0.2% Ficoll 400, 0.2% bovine serum albumin (BSA), 0.2% polyvinylpyrrolidone, 50 mM sodium phosphate pH 6.7, 100 µg ml⁻¹ denatured salmon-sperm DNA and 50% formamide. The samples were hybridized to four recombinant DNA probes physically homologous to loci unique to the subregion p15 of the short arm of human chromosome 11 (ref. 34), reveal restriction fragment length polymorphisms and exhibit close genetic linkage³¹. The probe pHins-310, homologous to the insulin gene³⁵, and the probe pTBB-2, homologous to the c-Ha-ras-1 gene³⁶, exhibit polymorphic alleles of various lengths, generated by insertions or deletions of DNA sequences, in *PvuII*- and *TaqI*-digested DNA, respectively³⁴⁻³⁶. The probe JW151 is homologous to the β -globin gene cluster³⁷ and reveals restriction fragment length alleles of the γ^G portion of the locus of 8.0 and 7.2 kilobases (kb) and alleles of 3.5 and 2.7 kb at the γ^A portion of the locus in *HindIII*-digested DNA³⁷. The probe pADJ-762, an arbitrary genomic DNA segment, reveals alleles of 2.3 and 1.8 kb in *MspI*-digested DNA and of 11.6 and 4.2 kb, as well as a constant 6.9 kb band, in *BclI*-digested DNA³³. The ³²P-labelled nick-translated probes were hybridized to the membranes at 47 °C for 24 h in 0.75 M NaCl/0.075 M sodium citrate, 0.02% Ficoll 400, 0.02% BSA, 0.02% polyvinylpyrrolidone, 20 mM sodium phosphate pH 6.7, 100 µg ml⁻¹ denatured salmon-sperm DNA, 10% dextran sulphate and 50% formamide. The membranes were washed once at room temperature in 0.3 M NaCl/0.03 M sodium citrate/0.1% SDS for 30 min and then twice for 30 min each at 65 °C in 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS. The membranes were then exposed to Kodak XAR-5 X-ray film backed with Dupont lightning-plus intensifying screens for 1-3 days. Alleles at each locus are designated 1 or 2 according to decreasing length; C indicates a common, invariant band. Numbers to the right of each autoradiograph indicate the molecular size of the indicated allele in kilobases, derived from the standards run with each gel.



could be revealed by mitotic segregation events, similar to those demonstrated previously for Wilms' tumour¹⁸⁻²¹ and retinoblastoma²², which would serve specifically to produce hepatoblastoma, rhabdomyosarcoma and adrenal carcinoma tumours that have lost constitutional heterozygosity. In particular, if such homozygosity was specifically obtained for chromosome 11p, the data would imply that the Wilms' tumour locus plays a pleiotropic role in developing embryonic kidney, liver, adrenal and muscle tissue, perhaps by differential stage-specific expression. Such data would also indicate a limited number of genetic loci in which recessive mutations play a role in tumour development, each with specificity for a larger subset of tissue types. To distinguish between these hypotheses, we performed studies similar to those described previously for retinoblastoma²² and Wilms' tumour¹⁸⁻²¹.

Hepatoblastoma development

Figure 1 displays the constitutional and tumour genotypes at loci on chromosome 11p for two hepatoblastoma patients. CHMC 429572 (Patient A) was diagnosed with embryonal hepatoblastoma at 6 years of age. He was constitutionally heterozygous (that is, displayed both parental fragment length alleles) at the two loci defined by pADJ-762 in *MspI*-digested DNA and pTBB-2 in *TaqI*-digested DNA. One of the alleles at each of these loci was missing in the primary hepatoblastoma biopsy tissue and in the first passage of the biopsy in an immunodeficient mouse (Fig. 1a). CHMC 410212 (Patient B) was diagnosed with embryonal hepatoblastoma at 1.2 years of age. She was constitutionally heterozygous at the three loci on

human chromosome 11p defined by pADJ-762 in *BclI*-digested DNA, pHins-310 in *PvuII*-digested DNA and the γ^A portion of the β -globin locus defined by JW151 in *HindIII*-digested DNA. Each of these loci had become homozygous or hemizygous in tissue obtained as a primary hepatoblastoma biopsy, in the primary tumour tissue transplanted twice in immunodeficient mice and in metastatic tumours from the right and left lungs of the same patient (Fig. 1b). A third patient, CHMC 304772, was diagnosed with hepatoblastoma with elements of hepatocellular carcinoma at 3.1 years of age. He was constitutionally heterozygous at the two loci defined by pADJ-762 in *MspI*-digested DNA and pTBB-2 in *TaqI*-digested DNA; both alleles at each of these loci were apparent in the tumour (data not shown).

These results are similar to those reported for Wilms' tumour¹⁸⁻²¹. The somatic loss of constitutional heterozygosity in the two hepatoblastoma tumours shown in Fig. 1 could be effected by abnormal mitotic events. Mechanisms that have been demonstrated include the non-disjunctional loss of the wild-type chromosome with reduplication of the mutant chromosome, non-disjunctional loss of the entire wild-type chromosome and deletion or loss of the region of the chromosome containing the marker loci¹⁸⁻²². To distinguish between these possibilities, DNA samples from the normal tissues and primary tumour biopsies described in Fig. 1 were hybridized concurrently with a DNA probe homologous to loci on chromosome 11 (pADJ-762) and chromosome 17 (DOSLC4) (ref. 24). The resultant hybridization signals were analysed by quantitative densitometry (Table 1). The peak area obtained with each signal was calculated and the ratios of the signal unique to chromosome 11 to that unique to chromosome 17 were compared between the normal and tumour

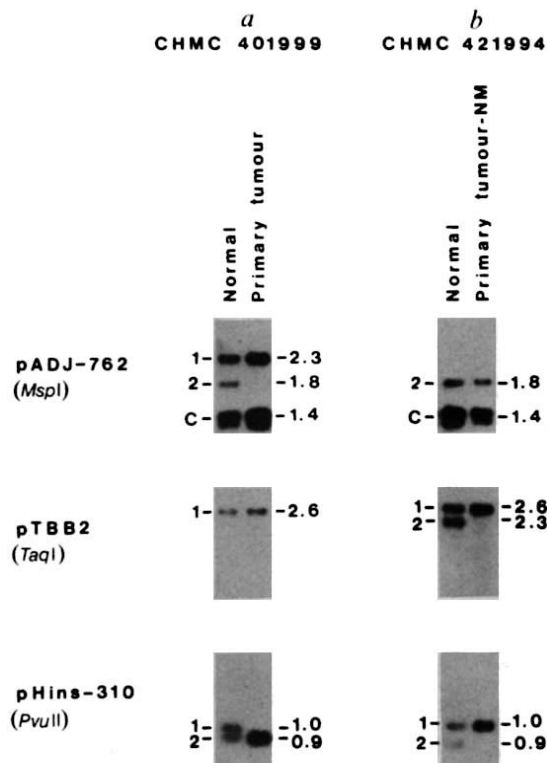


Fig. 2 Loss of constitutional heterozygosity at loci on human chromosome 11p in rhabdomyosarcoma; *a*, patient A; *b*, patient B. **Methods.** *a*, Normal-tissue DNA was isolated from peripheral leukocytes and tumour DNA isolated directly from the primary tumour specimen obtained at surgery. *b*, Normal-tissue DNA was isolated from peripheral leukocytes; tumour DNA was isolated from a first passage xenograft of the primary tumour (NM). The conditions for digestion of the DNA samples, electrophoresis, Southern transfer to nylon membranes, hybridization and autoradiography were as described previously^{18,22,26} and are detailed in Fig. 1 legend. Alleles at each locus are designated 1 or 2 according to decreasing length; C indicates a common, invariant band. Numbers to the right of each autoradiograph indicate the molecular size of the indicated allele in kilobases, derived from standards run with each gel.

tissues. In each case, these ratios were similar in the two tissues, indicating that the loci on chromosomes 11 and 17 were present in both normal and tumour tissue in the same copy number. These data indicate that the tumour specific losses of constitutional heterozygosity described for the hepatoblastomas in Fig. 1 were effected by non-disjunctional loss of the wild-type chromosome with reduplication of the mutant chromosome rather than by chromosomal loss or deletion. It is also possible that loss of heterozygosity at the loci tested here, all of which map to the telomeric portion of chromosome 11, could arise by mitotic recombination between the centromere and the marker loci. We are not presently able to distinguish between the recombination and loss/reduplication mechanisms; such distinctions await the development of probes homologous with loci between the mutant hepatoblastoma locus and the centromere.

Rhabdomyosarcoma development

Figure 2 displays constitutional and tumour genotypes of two rhabdomyosarcoma patients. CHMC 401999 (Patient A) was diagnosed with embryonal rhabdomyosarcoma at 0.8 years of age. She was constitutionally heterozygous at the two loci defined by pADJ-762 in *MspI*-digested DNA and pHins-310 in *PvuII*-digested DNA. CHMC 421994 (Patient B) was diagnosed with embryonal rhabdomyosarcoma at 1.1 years of age. She was constitutionally heterozygous at the loci homologous to pTBB-2 in *TaqI*-digested DNA and pHins-310 in *PvuII*-digested DNA.

Table 1 Densitometric dosage analysis of DNA from normal leukocytes and primary tumour biopsies from hepatoblastoma and rhabdomyosarcoma patients

Patient	Tissue	Chromosome 11 chromosome 17	Tumour normal
CHMC 429572	Leukocytes	2.84	1.12
	Hepatoblastoma	3.20	
CHMC 410212	Leukocytes	1.16	1.03
	Hepatoblastoma	1.20	
CHMC 401999	Leukocytes	2.24	1.06
	Rhabdomyosarcoma	2.38	
CHMC 421994	Leukocytes	3.15	0.93
	Rhabdomyosarcoma	2.92	

Normal and tumour DNA from the patients described in Figs 1 and 2 were digested with the restriction endonuclease *BclI*, transferred to nylon membranes and hybridized simultaneously to the radiolabelled probes pADJ-762 and DOSLC4, as described in Fig. 1 legend. These probes are homologous to unique loci on human chromosomes 11 (ref. 33) and 17 (ref. 24), respectively. The resultant autoradiogram was analysed by scanning densitometry and the peak areas corresponding to each hybridization signal were calculated by electronic integration. The peak areas specific to chromosomes 11 and 17 were compared in each sample and these normalized values were compared for each tumour/normal tissue pair.

Only one of the two codominant alleles at each of these loci was apparent in the primary tumour biopsy from Patient A and in the first-passage xenograft of the primary rhabdomyosarcoma from Patient B. Patient CHMC 405680 was diagnosed with alveolar rhabdomyosarcoma at 3.3 years of age. He was constitutionally heterozygous at the two loci on chromosome 11p15 defined by pADJ-762 in *BclI*-digested DNA and pTBB-2 in *TaqI*-digested DNA; heterozygosity was maintained at both loci in the corresponding tumour sample (data not shown). DNA isolated from normal and tumour samples from patients A and B was hybridized to DNA probes homologous to loci on chromosome 11 (pADJ-762) or chromosome 17 (DOSLC4). Densitometric analysis of the resultant autoradiographic signals, as described above for the hepatoblastoma cases, is shown in Table 1. Again, the number of copies of the two loci was similar in the normal and tumour tissues, indicating that the loss of constitutional heterozygosity in the tumours was effected by a mitotic mechanism which resulted in loss of the wild-type chromosome 11 homologue with maintenance and reduplication of the mutant chromosome, rather than by simple chromosomal loss or deletion of the segment of the chromosome containing the marker loci.

Chromosomal specificity

The somatic losses of constitutional heterozygosity described above for hepatoblastoma and rhabdomyosarcoma, and previously for Wilms' tumour¹⁸⁻²¹, could serve specifically to unmask initial predisposing mutations on chromosome 11 as shown previously for chromosome 13 in retinoblastoma^{22,23}. Alternatively, the attainment of chromosome 11 homozygosity in these tumours might represent a portion of a generalized loss of heterozygosity, such as that proposed based on analyses of isozymic alleles²⁵. To distinguish between these possibilities, the Southern transfers shown in Figs 1 and 2 were treated with alkali to remove the chromosome 11 probes²⁶ and then hybridized to several probes that reveal restriction site polymorphism on other human autosomes^{24,26-28}. In every case where constitutional heterozygosity was present, both codominant alleles could be distinguished in the corresponding tumour tissue (Table 2). Thus, the tumour from patient CHMC 429572 maintained heterozygosity at loci mapped to chromosomes 13, 14 and 22; the tumour from patient CHMC 410212 for loci mapped to chromosomes 6 and 13; and the tumours from patients CHMC 401999 and CHMC 421994 for loci mapped to chromosomes 13 and 20. In total, these tumours, which had become homozygous at several loci on chromosome 11, retained heterozygosity at other loci on chromosomes 6, 13, 14, 20 and 22. These data indicate that the losses of constitutional heterozygosity for chromosome 11 in these tumours are specific events rather than part of a generalized loss during tumorigenesis. The maintenance

Table 2 Alleles present at loci on chromosomes other than 11 in normal and tumour tissue from hepatoblastoma and rhabdomyosarcoma cases

Probe	Enzyme	Chromosome	Alleles present											
			Hepatoblastoma						Rhabdomyosarcoma					
			CHMC 429572			CHMC 410212			CHMC 401999		CHMC 421994			
			N	T ¹	T ²	N	T ³	T ⁴	T ⁵	T ⁶	N	T ⁷	N	T ⁸
p3C7	<i>MspI</i>	6	1, 1	1, 1	1, 1	1, 2	1, 2	1, 2	1, 2	1, 2	1, 1	1, 1	2, 2	2, 2
pL2.56	<i>HindIII</i>	6	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	2, 2	2, 2	2, 2	2, 2
p7F12	<i>MspI</i>	13	ND	ND	ND	ND	ND	ND	ND	ND	1, 2	1, 2	1, 2	1, 2
	<i>TaqI</i>	13	1, 2	1, 2	1, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2
p1E8	<i>MspI</i>	13	1, 1	1, 1	1, 1	1, 2	1, 2	1, 2	1, 2	1, 2	2, 2	2, 2	2, 2	2, 2
p9A7	<i>HindIII</i>	13	1, 2	1, 2	1, 2	1, 2	1, 2	1, 2	1, 2	1, 2	1, 1	1, 1	1, 2	1, 2
pAW101	<i>EcoRI</i>	14	1, 2	1, 2	1, 2	2, 2	2, 2	2, 2	2, 2	2, 2	ND	ND	ND	ND
DOSLC2	<i>MspI</i>	20	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	1, 1	1, 2	1, 2	1, 2	1, 2
DOSLC8	<i>TaqI</i>	22	1, 2	1, 2	1, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2	2, 2

Tumour and normal DNA from hepatoblastoma patients CHMC 429572 and CHMC 410212, as well as rhabdomyosarcoma patients CHMC 401999 and CHMC 421994, were digested with appropriate restriction endonucleases, transferred to nylon membranes, hybridized to the indicated ³²P-labelled probes, washed and exposed to X-ray film as described in Fig. 1 legend. Each of the probes is homologous with a locus on the indicated chromosome^{24,26-28}. Tumour samples were: T¹, primary surgical biopsy; T², first passage xenograft of T¹; T³, primary surgical biopsy; T⁴, second passage xenograft of T³; T⁵, third passage xenograft of metastatic T³ excised from the right lung; T⁶, first passage xenograft of metastatic T³ excised from the left lung; T⁷, primary surgical biopsy; T⁸, first passage xenograft. In all cases, normal tissue was obtained as peripheral leukocytes. N, normal tissue; T, tumour tissue; ND, not determined.

Table 3 Alleles at loci on chromosome 11 in normal and tumour tissues from Ewing's sarcoma and osteogenic sarcoma patients

Patient		Alleles present					pTBB3
		pADJ-762		JW151			
		<i>MspI</i>	<i>BclI</i>	γ^G - <i>HindIII</i>	γ^A - <i>HindIII</i>	<i>TaqI</i>	
A Ewing's sarcoma	1 CHMC 319634	N	2, 2	1, 2	2, 2	1, 2	1, 2
		T	2, 2	1, 2	2, 2	1, 2	1, 2
	2 CHMC 414222	N	2, 2	1, 2	2, 2	1, 1	1, 2
		T	2, 2	1, 2	2, 2	1, 1	1, 2
	3 CHMC 434672	N	2, 2	1, 2	1, 2	1, 1	1, 1
		T	2, 2	1, 2	1, 2	1, 1	1, 1
B Osteogenic sarcoma	1 CHMC 438160	N	2, 2	ND	1, 1	1, 2	1, 2
		T	2, 2	ND	1, 1	1, 2	1, 2
	2 CHMC 428687	N	2, 3	ND	2, 2	1, 2	1, 2
		T	2, 3	ND	2, 2	1, 2	1, 2
	3 CHMC 434642	N	2, 2	ND	2, 2	1, 2	1, 1
		T	2, 2	ND	2, 2	1, 2	1, 1

Tumour and normal DNA from Ewing's sarcoma patients CHMC 319634, CHMC 414222 and CHMC 434672, as well as osteogenic sarcoma patients CHMC 438160, CHMC 428687 and CHMC 434642, were digested with the indicated restriction endonucleases, transferred to nylon membranes, hybridized to the indicated radiolabelled probes, washed and exposed to X-ray film as described in Fig. 1 legend. Tumour (T) samples were xenografts of primary surgical biopsies except in the case of CHMC 434642 where primary surgical biopsy material was obtained. Normal (N) tissue was obtained as peripheral blood leukocytes. ND, not determined.

of heterozygosity in these tumours for loci mapped to chromosome 13 is of particular interest because the same markers have previously allowed detection of retinoblastoma-specific losses of constitutional heterozygosity^{22,23}. The data described in Table 2 also provide assurance that the tumour-specific differences from constitutional genotypes shown in Figs 1 and 2 do not arise from simple mispairing of the samples. In fact, we have examined these samples at many more non-chromosome 11 loci than those shown in Table 2 and we have not observed any instance in which the homozygous alleles in each pair of tissues did not precisely match.

Tumour specificity

The previously described clinical data, as well as the specific loss of heterozygosity for loci on human chromosome 11 during the genesis of rhabdomyosarcoma and hepatoblastoma (described above) and Wilms' tumour¹⁸⁻²¹, suggest a pathogenetic mechanism shared by these three clinically associated tumour types. To determine whether these events are com-

mon among more than these three childhood cancers, we performed the analyses described in Table 3. We isolated DNA from normal and tumour samples obtained from three children each with Ewing's and osteogenic sarcoma. These DNA samples were then examined with the probes homologous to loci on chromosome 11, exactly as described in Figs 1 and 2. In each case, constitutional heterozygosity was maintained in these tumours. It has been reported previously²² that another pediatric tumour, retinoblastoma, maintained chromosome 11 heterozygosity in several tumour samples. These results suggest that the loss of constitutional heterozygosity in rhabdomyosarcoma, hepatoblastoma and Wilms' tumour is specific to chromosome 11 and also displays specificity in those childhood tumours in which it is a pathogenetic mechanism.

Discussion

Here, we have described molecular genetic evidence suggesting that human embryonal tumours reflect abnormalities in genes important to normal human organismal development. These

data unite two widely disparate classes of observation: the high-frequency clinical associations between rare tumours and congenital malformation syndromes³; and the specific somatic loss of constitutional heterozygosity in the oncogenesis of embryonal cancers¹⁸⁻²².

Our data suggest a common chromosomal pathogenetic mechanism in rhabdomyosarcomas and hepatoblastomas similar to that previously shown in Wilms' tumours¹⁸⁻²¹. These results imply a common aetiology in the genesis of these three embryonal tumours that involves pleiotropic expression of a mutant allele at the Wilms' tumour (*WAGR*) locus on human chromosome 11 band p13 (ref. 29). Chromosomal mechanisms that serve to unmask these recessive mutant alleles by elimination of the balancing wild-type homologue occurring at times specific to the development of kidney, liver or striated muscle could then result in an embryonal tumour specific to the developing tissue of the organ. Pathological observations showing that embryonal tumours maintain immature differentiation states imply that the product of the *WAGR* locus is necessary for the ordered progression of normal differentiation to occur. Failure to produce the gene product would then result in the persistence of precursor embryonal cells of each organ and thereby account for persistent renal blastema or nephroblastomatosis. The loss of function effected by aberrant chromosomal events at the various times in organismal development might explain the aberrant differentiation that is observed as heteroplasia^{14,15} in these tumours and that leads to the previously unexplainable presence of striated muscle precursors, rhabdomyoblasts, in embryonal tumours of the kidney (Wilms' tumour) and the liver (hepatoblastoma). Thus, several lines of clinical evidence suggest a pathogenetic association between these three embryonal tumours. It is not known whether this association results from differential expression of mutations in the same gene or represents overlapping deficiencies in a larger multigene complex. This distinction cannot be drawn until the gene(s) is isolated and analysed. Clearly, however, the genesis of each of these rare tumours involves the specific loss of constitutional heterozygosity for human chromosome 11 and these events occur in these clinically associated tumours and in none of the other embryonal tumours we have examined. Although we have not yet been able to examine adrenal carcinoma tissues, it is reasonable to propose, based on these findings and clinical observations^{5-8,11,12,17,18}, that this tumour also arises following the attainment of chromosome 11 homozygosity.

Most developmental anomaly syndromes associated with high risk for tumour development segregate in families with the characteristics of a single autosomal dominant mutation³. A genetic component in the aetiology of Beckwith-Wiedemann syndrome has been proposed based on its autosomal dominant mode of inheritance⁴ and by the observation of a constitutional karyotypic abnormality of human chromosome 11p (ref. 30) in

a few patients with the syndrome. These observations, taken together with the evidence reported here, are consistent with the idea that a genetic locus on the short arm of human chromosome 11 is involved in the determination of the phenotypic manifestations of this disorder, including the developmental and oncological malformations. The distinction between these two types of abnormalities may arise by differential requirements for the product of the normal locus at varying times of development. Alternatively, the developmental malformations may represent the heterozygote phenotype, whereas the various tumour types may arise only after the attainment of complete homozygosity for a mutant allele at the locus. Either of these possibilities implies that loci in which defects give rise to heritable groups of developmental disorders are also responsible for the oncological disorders associated with the syndrome.

These data also have broad implications for genetic mapping. They indicate that the number of genetic loci whose recessive mutant alleles predispose to cancer is more limited than might have been suspected based on previous studies¹⁸⁻²². Certainly, it does not now seem as if there is one such locus for each tumour type, but it is likely that these loci will have broader, but specific, tissue preference. Studies to identify loci similar to *Rb-1* and *WAGR* may now be designed more effectively by selecting tumour types based on their clinical associations. This approach has proved to be useful in this study and our preliminary studies show that retinoblastoma and osteosarcoma share both clinical association and a molecular pathogenetic mechanism. Conversely, the identification of which chromosome carries a recessive predisposing mutation by the kinds of comparisons of tumour and constitutional genotypes described here, in tumours which cluster with high frequency in the various developmental anomaly syndromes, may expedite the selection of genetic markers for pedigree analysis^{31,32}. For example, the identification of chromosomes that undergo specific losses of constitutional heterozygosity during oncogenesis of astrocytoma, medullary thyroid carcinoma and medulloblastoma may lead to localization of the genetic locus in their clinically associated autosomal dominant syndromes of neurofibromatosis, multiple endocrine neoplasia and basal cell nevus syndrome, respectively. These sorts of analyses may accelerate the development of recombinant DNA-based premorbid detection of these diseases.

We thank Drs M. Wigler, H. Kazazian, G. Bell, P. Pearson and R. White for kindly providing some of the molecular probes. We also thank Drs K. Bove, J. McAdams, J. Noseworthy, J. Neely and P. Dignan for their advice and interest; Marianne Brown and Erika Simon for technical assistance; and Juliann Hansen and Becky Howland for manuscript preparation. This work was supported in part by grants from the National Eye Institute (USPHS), the National Cancer Institute (USPHS) and the American Cancer Society.

Received 22 March; accepted 12 June 1985.

- Willis, R. A. *Pathology of Tumours* 4th edn 7-8 (Butterworth, London, 1967).
- Nicholson, G. W. *J. path. Bact.* **34**, 711-730 (1931).
- Warkany, J. *Congenital Malformations* 1199-1271 (Year Book, Chicago, 1971).
- Best, L. G. & Hoekstra, R. E. *Am. J. med. Genet.* **9**, 291-299 (1981).
- Sotelo-Avila, C. & Gooch, W. M. *Perspect. Pediatr. Path.* **3**, 255-272 (1976).
- Sotelo-Avila, C., Gonzalez-Crussi, F. & Fowler, J. W. *J. Pediatr.* **96**, 47-50 (1980).
- Muller, S., Gädner, H., Weber, B., Vogel, M. & Riehm, H. *Eur. J. Pediatr.* **127**, 219-226 (1978).
- Riedel, H. A. *Pediatrics* **10**, 19-27 (1952).
- Miller, R. W., Fraumeni, J. F. & Manning, M. D. *New Engl. J. Med.* **270**, 922-927 (1964).
- Tefft, M., Vawter, G. F. & Mitus, A. *Am. J. Roentg.* **103**, 800-822 (1968).
- Milman, D. H. & Grayzel, D. M. *Am. J. Dis. Child.* **81**, 408-420 (1951).
- Schweisgoth, O. *Solid Tumors in Children* (Wiley, New York, 1982).
- Potter, E. L. *Normal and Abnormal Development of the Kidney*, 263 (Year Book, Chicago, 1972).
- Wigger, H. J. *Hum. Path.* **7**, 613-623 (1976).
- Sheehan, H. L. *J. path. Bact.* **33**, 251-258 (1930).
- Wilkins, L. & Ravitch, M. M. *Pediatrics* **9**, 671-681 (1952).
- Pack, G. T. & Miller, T. R. *AMA Archs Surg.* **73**, 1060-1062 (1956).
- Koufos, A. *et al. Nature* **309**, 170-172 (1984).

- Orkin, S. H., Goldman, D. S. & Sallan, S. E. *Nature* **309**, 172-174 (1984).
- Reeve, A. E. *et al. Nature* **309**, 174-176 (1984).
- Fearon, E. R., Vogelstein, B. & Feinberg, A. P. *Nature* **309**, 176-178 (1984).
- Cavenee, W. K. *et al. Nature* **305**, 779-784 (1983).
- Cavenee, W. K. *et al. Science* **228**, 501-503 (1985).
- Barker, D., Schafer, M. & White, R. *Cell* **36**, 131-138 (1984).
- Dracopoli, N. C. & Fogh, J. J. *natn. Cancer Inst.* **70**, 83-87 (1983).
- Cavenee, W., Leach, R., Mohandas, T., Pearson, P. & White, R. *Am. J. hum. Genet.* **36**, 10-24 (1984).
- Wyman, A. R. & White, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6754-6758 (1980).
- Skolnick, M. H., Willard, H. F. & Menlove, L. A. *Cytogenet. Cell. Genet.* **37**, 210-273 (1984).
- Riccardi, V. M., Sujansky, E., Smith, A. C. & Francke, U. *Pediatrics* **61**, 604-610 (1978).
- Turleau, C. *et al. Hum. Genet.* **67**, 219-221 (1984).
- White, R. *et al. Nature* **313**, 101-105 (1985).
- Gusella, J. F. *et al. Science* **225**, 1320-1326 (1984).
- Barker, D., Holm, T. & White, R. *Am. J. hum. Genet.* **36**, 1159-1171 (1984).
- Gerald, P. S. & Grzeschik, K. H. *Cytogenet. Cell. Genet.* **37**, 103-126 (1984).
- Bell, G. I., Selby, M. J. & Rutter, W. J. *Nature* **295**, 31-35 (1982).
- Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404-409 (1982).
- Antonarakis, S. E., Boehm, C. D., Giardina, P. J. & Kazazian, H. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 137-141 (1982).

Crystallographically aligned metal–oxide composite made by reduction of a directionally solidified oxide–oxide eutectic

A. Revcolevschi & G. Dhalenne

Laboratoire de Chimie Appliquée, Université Paris-Sud,
Bâtiment 414, 91405 Orsay Cédex, France

Aligned metal–oxide composite materials have been grown from the melt by directional solidification at eutectic composition of several binary metal–oxide systems. The resulting microstructures consist of metallic fibres $\sim 1\ \mu\text{m}$ in diameter imbedded in an oxide matrix^{1,2}. The occurrence of such fibrous structures, as opposed to lamellar structures, corresponds to eutectic compositions for which the volume fraction of the minor component is $<30\%$ (ref. 3). The work reported here presents a different route to the fabrication of aligned metal–oxide composites, making the production of fine-scale lamellar microstructures possible. The method consists of the chemical reduction of lamellar oxide–oxide eutectic structures in conditions of temperature and partial pressure of oxygen in which only one of the oxides is transformed to the corresponding metal. We show here how the method can be successfully applied to the preparation of aligned Ni–ZrO₂ lamellar structures.

Directionally-solidified oxide–oxide eutectics have been extensively studied in the past two decades, experimentation being carried out in more than 20 binary systems⁴. Several of them were examined in our laboratory within the framework of general studies of the structure and properties of oxide interfaces^{5–7}; these systems were chosen with the object of creating interfaces between a sodium chloride type structure (NiO) and other phases of cubic symmetry. The systems which were examined include: NiO–CaO (sodium chloride); NiO–NiAl₂O₄ (spinel); NiO–calcium-stabilized cubic ZrO₂ (fluorite); NiO–Y₂O₃ (fluorite-related).

Well-aligned eutectic microstructures were obtained and growth directions of the oriented structures, crystallographic relations existing between the two phases of each eutectic as well as the indices of the interface planes were determined^{8–10}.

It is of interest to extend these studies to metal–oxide interfaces by trying to reduce one of the oxides of an oxide–oxide eutectic structure into the corresponding metal. We chose the binary system NiO with stabilized cubic ZrO₂ because of the stability of ZrO₂ in conditions in which NiO could be easily

reduced into nickel. Another reason for this choice was the possible fabrication of a fine-scale structure made of alternate lamellae of ionically conducting (ZrO₂) and electronically conducting (Ni) phases.

Directional-solidification experiments in the NiO–ZrO₂ system were carried out in a floating-zone device¹¹. The starting rods which were used had a composition of 70 mol.% NiO, 25.5 mol.% ZrO₂, 4.5 mol.% CaO and were prepared by pressing (1 kbar) a mixture of finely ground 99.999% pure NiO, 99.9% pure ZrO₂ and 99.9% pure CaCO₃. The pressed rods, 12 mm in diameter, were sintered in air at 1,200 °C for 24 h before directional solidification was undertaken.

Growth experiments were performed in air. Optical examination of the microstructures indicate well aligned eutectic lamellar structures, the two phases of the eutectic being pure nickel oxide ($a_{\text{NiO}} = 4.178\ \text{\AA}$) and calcia stabilized cubic zirconia ($a = 5.13\ \text{\AA}$). The variation of the interlamellar spacing λ with the solidification rate R follows the law $\lambda^2 R = \text{constant}$, proposed by Tiller¹² and verified in many eutectic systems. Typically, for a growth rate of $1.5\ \text{cm h}^{-1}$, the interlamellar spacing is $\sim 3\ \mu\text{m}$ (Fig. 1).

Microfragments of the directionally solidified samples oriented along the growth direction were studied by rotating crystal, Weissenberg and Buerger precession methods, to determine preferred growth directions and epitaxial relations. Growth directions are mainly [110] for nickel oxide and [001] for calcia-stabilized zirconia: epitaxial relationships are $[1\bar{1}0]//[001]$ and $(111)//(100)$, the interface planes being $(111)_{\text{NiO}}$ and $(100)_{\text{ZrO}_2}$ (ref. 13).

The information relative to growth directions of the eutectic phases was confirmed by X-ray diffractometry experiments carried out on transverse sections of the samples: the only peaks observed were high-intensity peaks corresponding to planes normal to growth directions, that is (220) for NiO and (200) and (400) for calcia-stabilized zirconia.

Directionally solidified samples, 3-mm thick, were submitted for 70 h to a chemical reduction treatment at 1,075 °C under a CO/CO₂ mixture (CO/CO₂ = 0.1) corresponding to $P_{\text{O}_2} = 1.6 \times 10^{-11}$ atmospheres, that is, conditions where NiO should be totally reduced to nickel¹⁴.

The microstructure resulting from reduction of the samples presented in Fig. 2 indicates that the aligned nature of the original eutectic structure and the cohesion between the two phases are maintained. Powder X-ray diffraction experiments on crushed samples show total reduction of NiO, the only observable lines being those of pure nickel ($a = 3.52\ \text{\AA}$) and of unaltered calcia-stabilized zirconia, the lattice parameter of which is the same as that found in the eutectic structure.

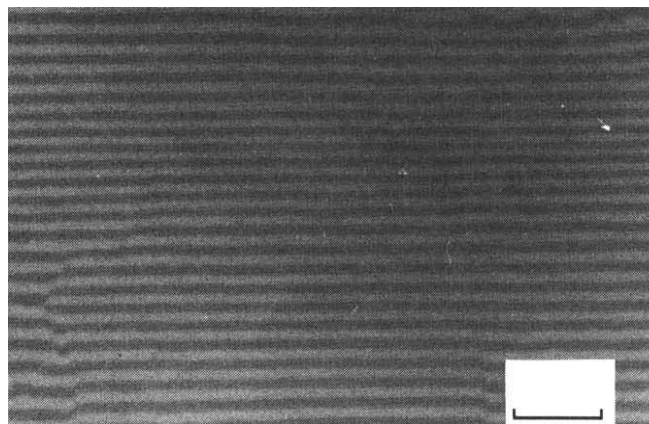


Fig. 1 Transverse section of the nickel oxide–calcium oxide–stabilized zirconia eutectic (70 mol.% NiO, 30% ZrO₂–CaO) solidified at $1.5\ \text{cm h}^{-1}$. Scale bar, $10\ \mu\text{m}$.

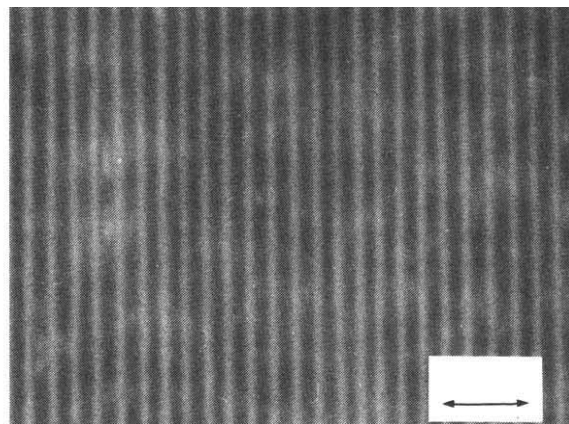


Fig. 2 Crystallographically aligned metal–oxide composite structure made by reduction of a directionally solidified oxide–oxide eutectic. Scale bar, $10\ \mu\text{m}$.

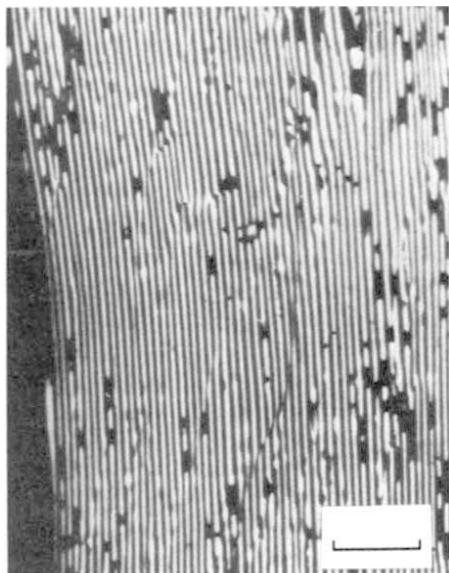


Fig. 3 Characteristic feature of a Ni-ZrO₂ sample showing presence of elongated cavities in the Ni phase. A crack can be seen on the left. Scale bar, 30 μm.

In some areas of the samples, elongated cavities presumably caused by the elimination of oxygen from the nickel oxide phase can be observed in the nickel phase. A few cracks also occur, the region adjacent to these being depleted in cavities (Fig. 3), indicating cavity coalescence. Observation at higher magnification by scanning electron microscopy confirms these features and particularly the boundary coherence. A change in lamellar spacing on reduction was not detectable. This can be explained by the fact that despite the volume change when transforming from NiO to Ni, the resulting modification of lamellar spacing is theoretically small enough to lie within experimental error of the measurements.

X-ray diffractometry experiments performed on transverse sections of samples of oxide-oxide eutectic structure submitted to reduction by CO/CO₂ mixtures show, compared with the powder diffractograms of the same material, very strong (220) nickel and (200) and (400) zirconia peaks (Fig. 4).

Considering the above results concerning the orientation of the two phases of the NiO-ZrO₂ samples, it seems that the oriented nature of the ZrO₂ phase remains unchanged; the nickel phase seems also to be perfectly oriented, the (110) orientation of the nickel lamellae being that of the original nickel oxide phase. Taking into account the f.c.c. structure of nickel and the NaCl-type structure of nickel oxide, it appears that the reduction takes place by a topotactical process in which the departure of oxygen from the NiO phase leaves intact the f.c.c. nickel sublattice which contracts, preserving the monocrystalline nature of the nickel-based phase. This mechanism was confirmed by separate X-ray diffractometry experiments in which large monocrystalline NiO samples which were chemically reduced in the conditions described above showed transformation into mostly monocrystalline nickel.

These results may be compared with those obtained by Little *et al.*¹⁵ who examined by transmission electron microscopy (TEM) the early stages of low-temperature (200 °C) reduction of nickel oxide and observed the formation of both epitaxial and non-epitaxial nickel nuclei, the latter being limited in number at the beginning of the process but becoming predominant as reduction time progressed. Little *et al.* concluded that loss of epitaxy is a feature of the growth of the nuclei. The presence in our experiments of non-epitaxial nickel, if any, was minor, as indicated by Fig. 4b. One can rationalize the difference in behaviour observed in the two sets of experiments by the large temperature gap which separates them. At higher temperature, more nucleation would lead to small epitaxial nuclei which would rapidly coalesce into larger grains of quasi-iden-

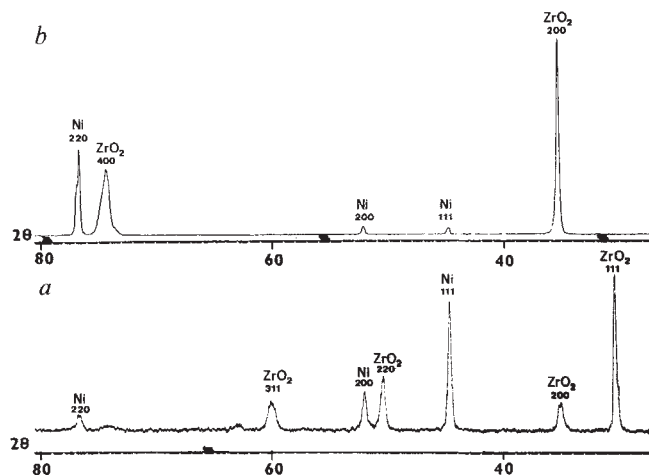


Fig. 4 X-ray diffractometry of the material resulting from reduction of eutectic NiO-ZrO₂: a, powder pattern; b, transverse section of eutectic structure.

tical orientation. An additional reason for the retention of the nickel oxide orientation in nickel may be found in the crystallographic constraint provided by the ZrO₂ phase.

We have shown, therefore, that reduction of lamellar oxide-oxide eutectics may lead to fine-scale alternate metal-oxide structures. The preliminary X-ray studies indicate that, at least in the system under examination, the crystallographic relationships between the two phases of the original structure remain unchanged after reduction of one of the oxides. More elaborate X-ray and TEM studies are being performed to identify finally the Miller indices of the interface planes which apparently also remain unchanged, and second to better describe the Ni-ZrO₂ interface.

Received 6 February; accepted 2 May 1985.

1. Briggs, J. & Hart, P. E., *J. Am. ceram. Soc.* **59**, 530-531 (1976).
2. Ashbrook, R. L. *J. Am. ceram. Soc.* **60**, 428-435 (1977).
3. Hunt, J. D. & Jackson, K. A. *Trans. AIME* **236**, 843-852 (1966).
4. Stubican, V. S. & Bradt, R. C. *Ann. Rev. mater. Sci.* **11**, 267-297 (1981).
5. Dhalenne, G., Revcolevschi, A. & Gervais, A. *Phys. Status Solidi A* **56**, 1, 267-276 (1979).
6. Dhalenne, G., Revcolevschi, A. & Monty, C. *Phys. Status Solidi A* **56**, 1, 623-636 (1979).
7. Dhalenne, G., Déchamps, M. & Revcolevschi, A. *J. Am. ceram. Soc.* **65**, C11-C12 (1982).
8. Fragneau, M. & Revcolevschi, A. *J. Am. ceram. Soc.* **66**, C162-C163 (1983).
9. Fragneau, M., Revcolevschi, A. & Michel, D. *Adv. Ceram.* **6**, 110-116 (1983).
10. Dhalenne, G. & Revcolevschi, A. *J. Cryst. Growth* **69**, 616-618 (1984).
11. Revcolevschi, A. *Rev. Int. Hautes Temp. Réfract.* **7**, 73-90 (1970).
12. Tiller, W. A. in *Liquid Metals and Solidification*, 276 (American Society of Metals, Cleveland, 1958).
13. Dhalenne, G., d'Yvoire, F. & Revcolevschi, A. *J. Physique* **46** (C4) 441-447 (1985).
14. Dominguez-Rodriguez, A. & Castaing, J. *Rev. Phys. Appl.* **11**, 387-391 (1976).
15. Little, J. A., Evans, J. W. & Westmacott, K. H. *Met. Trans.* **B11**, 519-524 (1980).

Continental volume and freeboard through geological time

G. Schubert & A. P. S. Reymer

Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

Approximate constancy of continental freeboard since the Archaean¹⁻³ has been used to argue that no growth of the continental crust occurred during the Proterozoic and Phanerozoic^{4,5} and that there was limited vertical movement of old cratons.⁶ However, it has been shown^{7,8} that constancy of freeboard implies net growth of the continental crust, because the ocean basins deepen as the heat flow from the mantle decreases with time (crustal thickness d_c , the volume of the oceans V_o and the Earth's surface area A_e being constant). A net growth of about 25% or 1 km³ yr⁻¹ is necessary to maintain freeboard at a constant value⁷. We further explore here the consequences of approximate constant freeboard for continental growth using a model that relates the volumes of

isostatically compensated continents and oceans to the secular decline in terrestrial heat flow. We find that a post-Archaeon increase in freeboard by 200 m requires continental growth of only 10%, while a decrease in freeboard by 200 m during this same period necessitates a crustal growth of 40%. Shrinkage of the continental crust since the end of the Archaeon can be ruled out. Changes of >10% in post-Archaeon crustal thickness are highly unlikely.

Temporal variations in freeboard (transgressions and regressions of sea level) can arise from melting of polar ice, post-glacial rebound, changes in the number and areal distribution of continental segments⁹, and changes in the volume of mid-ocean ridges and continents. We consider only the secular changes in freeboard due to changes in the area and thickness of the continents, and changes in the volume of the oceans particularly as a consequence of the deepening of the ocean basins with time on a cooling Earth.

There is an uncertainty of a few hundred meters in the difference between freeboard during the early Proterozoic and the present, mainly due to our inability to calibrate the significance or representativeness of present-day freeboard. The uncertainty in how freeboard has changed since the Archaeon, although small, has important consequences for crustal growth.

Continental crustal volume V_c is the product of its area A_c and thickness d_c . Ocean volume V_o is given by $V_o = A_o(d_a + d_b)$, where A_o is the area of the oceans, d_b is the mean depth below the level of the mid-ocean ridges (MOR) and d_a is the depth below sea level of the rise crests (Fig. 1). The areas of the oceans

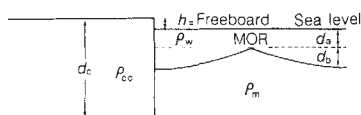


Fig. 1 Sketch of the model. MOR, Mid-ocean ridge; ρ_m , mantle density; ρ_{cc} , crustal density; ρ_w , water density; h , continental freeboard.

and continents sum to the total surface area of the Earth. We assume that A_c has remained constant throughout geological time (the change in A_c due to whole Earth cooling is negligible); V_o is also taken to be constant since the end of the Archaeon. This assumes that all the water in the present oceans outgassed from the interior at the earliest stages of the Earth's evolution and that the amount of water in the form of ice and vapour was not substantially different 2,500-Myr ago from the amount at present.

Continents and oceans are assumed to have always been in isostatic balance. Application of the principle of isostasy¹⁰ to the model of Fig. 1 gives (see Fig. 1 for parameter definitions)

$$d_a - d_a^* = (d_c - d_c^*) \frac{(\rho_m - \rho_{cc})}{(\rho_m - \rho_w)} - (h - h^*) \frac{(\rho_m)}{(\rho_m - \rho_w)} \quad (1)$$

An asterisk indicates present-day values. Ocean-continent density differences extending to great depth in the mantle and varying through geological time could modify the isostatic balance. The connection¹⁰ between d_b and the mean oceanic surface heat flow q_s is

$$\frac{d_b}{d_b^*} = \frac{q_s^*}{q_s} \quad (2)$$

It is generally accepted that q_s has decreased monotonically through geological time through the gradual decline in radiogenic heating and the secular cooling of the Earth¹¹. The decrease tends to follow the decline in internal radioactive heat sources¹¹ and we have adopted a model¹⁰ of the decrease in heat production to derive the form of $q_s(t)$ shown in Fig. 2. Mean oceanic heat flow decreases by about a factor of 4 since 4,500 Myr ago; the decline since the end of the Archaeon is by a factor of ~2. We have ignored the small amount of mantle heat that escapes through the continents consistent with our simplified model and

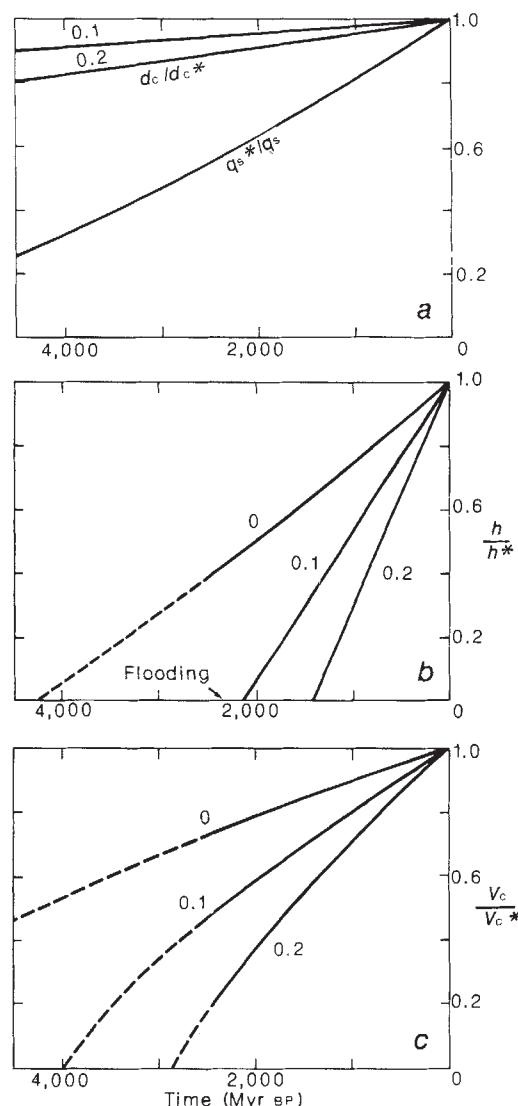


Fig. 2 a, Assumed variations of crustal thickness and mean oceanic heat flow through geological time. b, Continental volume (as a percentage of present volume V_c^*) versus time before present for different assumed crustal thickness variations. c, Freeboard (as a percentage of its present value h^*) as a function of time for V_c and d_c constant. The parameter labelling the curves in b and c is the fractional increase in crustal thickness over 4,500 Myr. The dashed portions of the curves refer to the time period in Earth history when freeboard and crustal thickness are highly uncertain. If h is negative, the continents are flooded. Thus the intersection of the curves in c with the horizontal axis gives the time of emergence of the continents above sea level.

the large uncertainty in past surface heat flow. Figure 2 also shows two models of crustal thickening that we consider; the values 0.1 and 0.2 refer to thickness increases of 10 and 20% over geological time.

Our model relating crustal growth and continental freeboard is similar in many respects to the one studied by Wise¹. Our principal contribution is the addition of equation (2) and the recognition that the cooling of the Earth deepens the ocean basins and thereby has a major role in controlling freeboard.

For constant freeboard and crustal thickness

$$\frac{V_c}{V_c^*} = \frac{A_c}{A_c^*} \frac{(V_o/A_c^*)}{d_a^* + d_b^*(q_s^*/q_s)} \quad (3)$$

in agreement with our previous result⁷. The increase of V_c with time as a consequence of the decline in q_s is readily deduced from equation (3); $V_c(t)$ is shown in Fig. 3. Parameter values are summarized in Table 1; q_s^*/q_s is given in Fig. 2a. The curve labelled 0 in Fig. 2b is for d_c constant and shows that a 25%

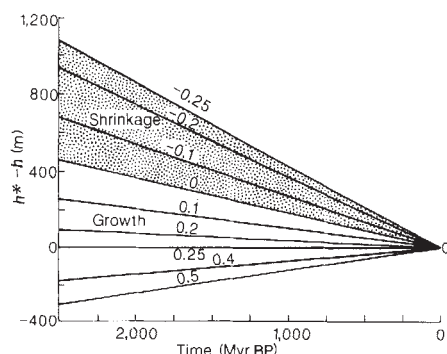


Fig. 3 Freeboard variation with time during the Proterozoic-Phanerozoic for prescribed percentage changes in continental volume between 2,500 Myr ago and the present. Both crustal growth and shrinkage are considered. Crustal thickness is assumed to be constant.

post-Archaeon increase in crustal volume is necessary to maintain constant freeboard⁷.

Since constant freeboard requires 25% post-Archaeon crustal growth (with d_c constant), it is important to determine what change in freeboard would occur for V_c = constant. With both V_c and d_c constant,

$$h^* - h = \left(1 - \frac{\rho_w}{\rho_m}\right) \left(1 - \frac{q_s^*}{q_s}\right) d_b^* \quad (4)$$

Equation (4) shows that freeboard must have been smaller in the past. The decrease in q_s and the deepening of the ocean basins causes a secular regression in sea level, that is, an increase in freeboard. Figure 2c illustrates the variation in freeboard with time computed from equation (4) and the parameter values given in Table 1. At the end of the Archaeon, h is predicted to be only about 40% of its present value (curve 0); freeboard must have increased by 400 m. This seems inconsistent with the geological constraint of approximate constant freeboard, but it is difficult to estimate the uncertainty in this constraint quantitatively.

It is clear from the above that freeboard variations of only several hundred metres are associated with crustal volume variations of tens of per cent. Since freeboard changes of this magnitude cannot be ruled out on geological grounds we consider the general relation between h and V_c for constant d_c .

$$h - h^* = \frac{(\rho_m - \rho_w)}{(\rho_m)} \left[d_a^* + d_b^* \frac{q_s^*}{q_s} - \frac{V_0}{(A_c - V_c/d_c^*)} \right] \quad (5)$$

We have evaluated equation (5) for prescribed changes in continental volume between the end of the Archaeon and the present and show the results in Fig. 3.

As we have already seen, no change in freeboard implies crustal growth of 25% during the post-Archaeon. If freeboard has been within ± 200 m of its present value since the end of the Archaeon, then the continents have grown during the past 2,500 Myr by anywhere from ~ 10 to 40%. Crustal growth is required as long as freeboard at the end of the Archaeon was not more than 400 m smaller than it is today. Crustal shrinkage since the end of the Archaeon requires an increase in freeboard by at least ~ 0.5 km since 2,500 Myr ago. Although secular freeboard changes of several hundred metres since the end of the Archaeon might be consistent with the geological evidence, changes of 0.5 km or more probably can be ruled out. Therefore, we conclude on the basis of Fig. 3 that post-Archaeon crustal growth by a few tens of per cent is required.

Cooling of the Earth and deepening of the ocean basins tend to increase freeboard through geological time. Crustal shrinkage (with constant thickness) reinforces this trend. Figure 3 shows that shrinkage by even as much as 10% since 2,500 Myr ago would mean that the continents just emerged above sea level at the end of the Archaeon. Crustal growth offsets the tendency of deepening ocean basins to increase freeboard with time. If the crust has grown more than 25% since the end of the

Table 1 Values of model parameters

Parameter	Definition	Value
V_0	Volume of oceans	$1.17 \times 10^{18} \text{ m}^3$
V_c^*	Present volume of continents	$7.6 \times 10^{18} \text{ m}^3$
A_c^*	Present area of continents	$2.17 \times 10^{14} \text{ m}^2$
A_0^*	Present area of oceans	$2.93 \times 10^{14} \text{ m}^2$
A_e	Surface area of Earth	$5.1 \times 10^{14} \text{ m}^2$
d_a^*	Present depth of mid-ocean ridge system below sea level	2,500 m
d_b^*	Present depth of ocean basins below the rise crests	1,500 m
d_c^*	Present average crustal thickness	38 km
h^*	Present continental freeboard	750 m
ρ_m	Mantle density	$3,300 \text{ kg m}^{-3}$
ρ_w	Water density	$1,000 \text{ kg m}^{-3}$
ρ_{cc}	Crustal density	$2,750 \text{ kg m}^{-3}$

Archaeon, freeboard has decreased during the Proterozoic-Phanerozoic, that is, the continents would have had their maximum emergence at 2,500 Myr ago. This runs against most assessments of the geological and sedimentological records, which support a constancy^{1,2,4-6} or a slight increase¹² rather than a decrease¹³ in continental freeboard since $\sim 2,500$ Myr ago.

To assess how variations in crustal thickness might affect our conclusions, we consider how freeboard and continental volume are related for prescribed $d_c(t)$

$$h - h^* = \frac{(\rho_m - \rho_w)}{(\rho_m)} \left[\frac{(\rho_m - \rho_{cc})}{(\rho_m - \rho_w)} (d_c - d_c^*) + d_a^* + d_b^* \frac{q_s^*}{q_s} - \frac{V_0}{(A_c - V_c/d_c)} \right] \quad (6)$$

With $d_c(t)$ from Fig. 2a, we use equation (6) to determine freeboard (or continental volume) given an assumed variation in V_c (or h). It can be deduced from equation (6) that crustal thickening augments the deepening of the ocean basins with time by requiring even larger increases in crustal volume (compared with d_c constant) to preserve constant freeboard.

Figure 2b illustrates how crustal thickening by 10 and 20% since 4,500 Myr ago modifies the amount of crustal growth required to maintain constant freeboard. Even a small amount of crustal thickening greatly increases the quantity of crustal growth necessary for constancy of freeboard. For example, with 10% crustal thickening over all of geological time, continental volume would have to have doubled since the end of the Archaeon. The geological evidence argues against such a large post-Archaeon increase in V_c .

Figure 2c summarizes how the assumed crustal thickening influences freeboard for no crustal growth. With V_c constant, a thinner crust in the past implies a larger continental area and, therefore, a smaller freeboard than at present. With even small increases in d_c , the continents would not rise above the oceans until at least the mid-Proterozoic. An increase in d_c and a late emergence of the continents have been proposed¹⁴⁻¹⁶, but this is not supported by geological data². A drastic global increase in d_c by underplating¹⁶ neglects the freeboard constraint and hotspot frequency data⁷, but underplating could be applicable to rapidly formed crustal segments¹⁷. We conclude that post-Archaeon changes in mean crustal thickness are highly unlikely.

Some secular changes in h may have occurred which are not included in our model, but their analyses are not useful as long as no firmer constraints can be put on the 'error' in constant freeboard. A regression during the Phanerozoic has been proposed¹², but the amount depends on an arbitrary point in time from which detailed data are available and does not necessarily point to a regression since the Archaeon¹⁸. A 5% change in ocean volume corresponds to a 200-m change in sea level. Therefore, small secular changes in ocean volume cannot be excluded. Present estimates of water subducted in trenches and released at mid-ocean ridges vary widely¹⁹ and need revision. Changes with time in the thickness of ocean floor sediments, in

the depth to ridge crests, in oceanic crustal thickness and in the development of deep mantle roots under the continents may all influence freeboard. Lack of data limits the usefulness of incorporating these effects into our model and increases the overall uncertainty in our results.

The continents increased their size during the Proterozoic and Phanerozoic^{7,17}, indicating at least 20% growth since 2,500 Myr ago. Such a growth rate is close to the requirement for constant freeboard and encourages us to believe that ocean volume and crustal thickness have remained essentially constant since 2,500 Myr ago, with error margins as given above.

We thank Donald U. Wise for constructive comments. This study was done under NASA grant NAG-9-76 while A.R. was on a sabbatical leave from the Hebrew University of Jerusalem.

Received 4 February; accepted 15 May 1985.

1. Wise, D. U. in *Geology of Continental Margins* (eds Burk, C. A. & Drake, C. L.) 45-58 (Springer, New York, 1974).
2. Windley, B. F. *Nature* **270**, 426-428 (1977).
3. Ambrose, J. W. *Am. J. Sci.* **262**, 817-857 (1964).
4. Armstrong, R. L. *Rev. Geophys. Space Phys.* **6**, 175-200 (1968).
5. Armstrong, R. L. *Phil. Trans. R. Soc. A* **301**, 443-472 (1981).
6. Watson, J. V. *Phil. Trans. R. Soc. A* **280**, 629-640 (1976).
7. Reymier, A. & Schubert, G. *Tectonics* **3**, 63-77 (1984).
8. Dewey, J. F. & Windley, B. F. *Phil. Trans. R. Soc. A* **301**, 189-206 (1981).
9. Harrison, C. G. A. *et al. Earth planet. Sci. Lett.* **54**, 1-16 (1981).
10. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
11. Schubert, G., Stevenson, D. & Cassen, P. *J. geophys. Res.* **85**, 2531-2538 (1980).
12. Hallam, A. *Nature* **269**, 769-772 (1977).
13. Abbott, D. *Tectonics* **3**, 709-722 (1984).
14. Hargraves, R. B. *Science* **193**, 363-371 (1976).
15. Hargraves, R. B. in *Precambrian Plate Tectonics* (ed. Kroner, A.) 21-56 (Elsevier, Amsterdam, 1981).
16. McKenzie, D. *Nature* **307**, 616-618 (1984).
17. Reymier, A. P. S. & Schubert, G. *Geology* (submitted).
18. McLennan, S. M. & Taylor, S. R. *Nature* **306**, 169-172 (1983).
19. DeVore, G. W. *Lithos* **16**, 255-263 (1983).

Sea-bed noises reveal role of turbulent bursting phenomenon in sediment transport by tidal currents

A. D. Heathershaw* & P. D. Thorne

Institute of Oceanographic Sciences, Crossway, Taunton TA1 2DW, Somerset, UK

In many engineering and sedimentological applications, it is usual to relate sediment movement to the time-averaged bed shear stress. However, the shear stress exerted on the sea bed by tidal currents is intermittent in nature due to the turbulent bursting phenomenon^{1,2}. Sediment movement at the sea bed also occurs in bursts^{3,4}. However, attempts to link these two processes have failed because of the poor temporal resolution of existing sediment-transport measurement techniques. Recent experimental advances now make this possible and we describe here the first detailed observations of sediment movement and fluid turbulence in tidal currents flowing over sandy gravels. These results show that high stress values are a necessary, but not sufficient, condition for bed-load sediment movement and that form drag on individual particles may have greater dynamical significance than the stress. This work also suggests that it may now be possible to describe sediment movement at a level which is deterministic in terms of the turbulence and at very much shorter timescales than have been used previously.

The importance of the bursting phenomenon in tidal current boundary layers has been confirmed in previous work^{1,2}. This has shown that, similarly to laboratory boundary layers⁵⁻⁸, the bulk of the Reynolds stress is made up by contributions from large and intermittent momentum transporting events which, collectively with other weaker motions, comprise the bursting phenomenon. Here the Reynolds stress is defined as $\tau = -\rho \overline{uw}$ where ρ is the fluid density, u and w are the horizontal and

vertical turbulent velocity fluctuations and the overbar denotes averaging in time. In particular, these early studies established that positive contributions to the stress were principally from two types of motion; ejections in which $u < 0$; $w > 0$ and sweeps in which $u > 0$; $w < 0$. In addition, two weaker motions, inward ($u < 0$; $w < 0$) and outward ($u > 0$; $w > 0$) interactions, made negative contributions to the stress. Note that the motions which contribute to the bursting sequence of events are generally considered to be associated with pairs of counter rotating vortices⁹. These, in turn, may be associated with a Taylor-Görtler instability¹⁰. However, the exact generation mechanism for the vortices and the bursting phenomenon is unknown⁹.

In parallel with the purely fluid dynamical interest, there has been considerable speculation¹¹⁻¹⁴ over the exact role of the turbulent bursting phenomenon in sediment transport. However, only recently has the development of a passive acoustic technique made it possible to observe sediment movement with a temporal resolution comparable with the high frequency turbulence in the flow. In particular, it is possible to detect¹⁵⁻¹⁷ the self-generated noise (SGN) arising from interparticle collisions of mobile sediment and to relate the acoustic intensity to the amount and size of sediment which is moving. Thus, for the first time, it has been possible to make a detailed examination of the relationship between bed-load transport and the turbulent structure of the boundary layer near the sea bed.

During studies to examine the incipient motion of sea-bed gravels in the West Solent, on the south coast of England⁴, measurements have been made of the horizontal (u) and vertical (w) turbulent velocity fluctuations at a height of 0.33 m above the sea bed using electromagnetic current meters¹⁵. The water depth in the study area was ~14 m. Underwater television (TV) and a wideband (-3 dB ~ 150 kHz) hydrophone were used to monitor the movement of sediments in the approximate size range 4-32 mm (ref. 4). The hydrophone was positioned at a height of 0.24 m above the sea bed. Turbulent flow and SGN data were recorded at 0.2-s intervals. (Further details of these measurements are given in refs 4, 15.)

Figure 1 shows measurements of the horizontal and vertical turbulent velocity fluctuations u and w , and their contribution $-uw$ to the Reynolds stress. Figure 1 confirms the presence of the bursting phenomenon in tidal currents flowing over gravels in the West Solent with instantaneous stress values which may be 10 or more times greater than the mean value obtained by averaging the stress over a typical record length of 1,200 s (Table 1). Figure 1 shows that the bursting phenomenon is uniquely identified by the large and intermittent contributions that it makes to the Reynolds stress. This intermittency occurs despite the fact that u and w are normally distributed and randomly fluctuating quantities when taken separately¹⁸. Figure 1 also shows the SGN signal, corrected for background noise, and illustrates that the bulk of the gravel movement occurs in bursts, during sweep type events. This is illustrated further in Fig. 2, which shows u , w , $-uw$ and the SGN signal on an expanded timescale. Figure 2a shows ejection type events ($u < 0$; $w > 0$) which produce no sediment movement while Fig. 2b shows sweep events ($u > 0$; $w < 0$) producing appreciable quantities of sediment movement. Underwater TV observations¹⁵ of gravel movement have provided an independent means of checking the performance of the SGN device as shown in Fig. 1. As far as we know this is the first description of observations of the type shown in Figs 1 and 2.

To examine in more detail the relationship between sediment movement and turbulent fluctuations in the Reynolds stress, comparisons have been made of the amounts of sediment movement occurring in the different types of Reynolds stress event. For these comparisons five records from the West Solent were chosen, corresponding to a period of decelerating tidal flow on 19 September 1982 (Table 1).

Acoustic intensity levels were compared with the stress ($-uw$) records, following linear trend removal from u and w , an empirically determined 1-s adjustment for lag effects due to the separation between the flow sensors and the sea bed, and possible

* Present address: Ocean Science Division, Admiralty Research Establishment, Portland, Dorset DT5 2JS, UK.

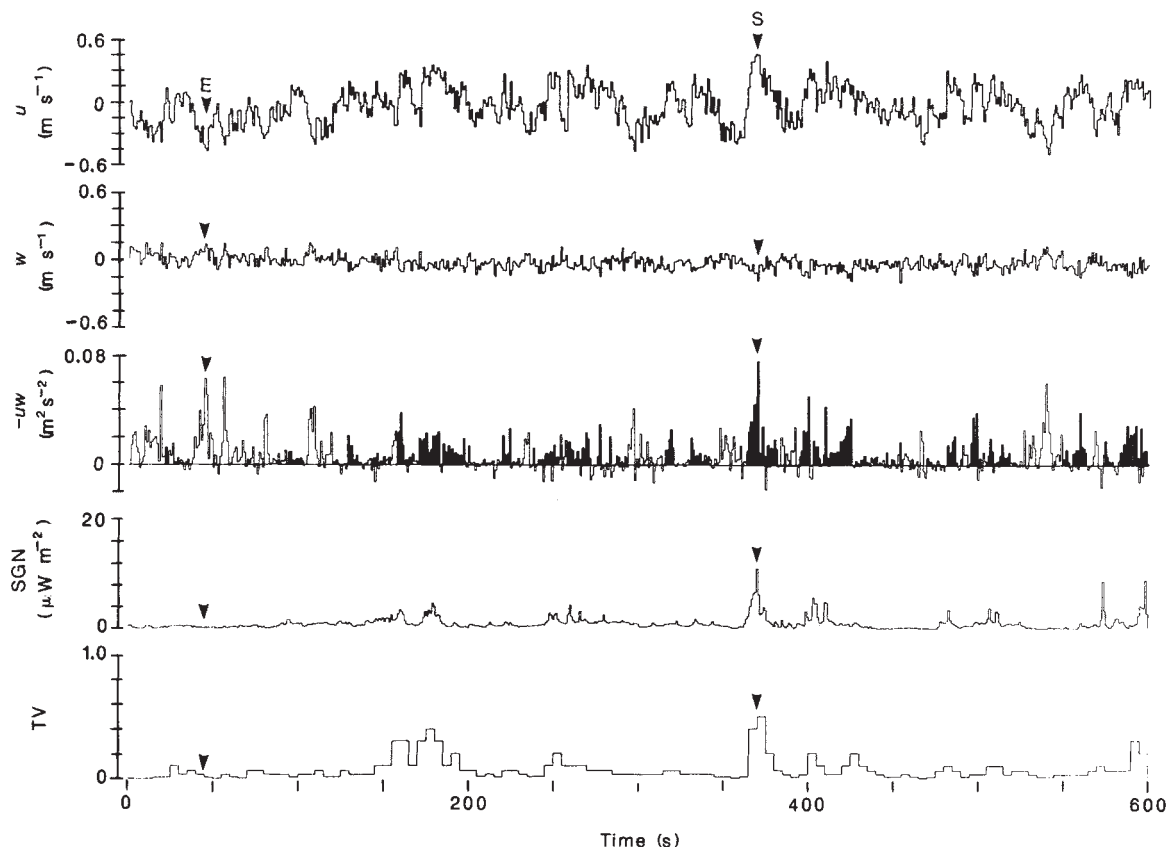


Fig. 1 Records of the horizontal and vertical turbulent velocity fluctuations u and w , and their product $-uw$. Also shown is the SGN acoustic intensity record and qualitative estimates of gravel movement from the underwater TV. For clarity 1-s averages only of u , w and SGN are shown together with 5-s estimates of gravel movement from the underwater TV. Black stress events (S) correspond to sweeps ($u > 0$; $w < 0$) and white events (E) correspond to ejections ($u < 0$; $w > 0$).

particle inertia. The significant stress events in each record were identified using a procedure similar to that given in refs 19, 20. Stress events were ranked according to peak size and regardless of sign, in such a way as to make up a total contribution of 90% to the overall stress. Event duration was defined by forwards and backwards interpolation to give uw values which were 10% of each peak. Event amplitude was defined as the average of $-uw$ taken over the duration of the event. Average values of u , w and the acoustic intensity were also calculated over the same time interval. Stress events were then identified as ejections, sweeps and inward or outward interactions according to the sign of the u and w averages. In this way, the average acoustic intensity levels, and hence the quantity of sediment moved by the significant stress events in each record, could be categorized and compared. Further details of the amplitudes, durations and percentage contributions to the overall stress, for the various events, are given in Table 1.

The results of this analysis are shown in Fig. 3, which shows

comparisons for 3,187 events measured over a period of ~100 min. This demonstrates quite clearly (Fig. 3a, b) that acoustic intensity levels associated with sweeps ($u > 0$; $w < 0$) are substantially higher than the intensity levels during ejections ($u < 0$; $w > 0$) at high stress values. This difference increases as the stress ($|uw|$) increases. From this we conclude that of the two types of motion which contribute to the bulk of the Reynolds stress, ejections and sweeps (Table 1), only sweeps are capable of supporting appreciable sediment movement. However, an interesting and unexpected result is that outward interaction events ($u > 0$; $w > 0$), although weaker and less frequent than sweeps (Fig. 3e), are capable of supporting greater sediment movement in the stress range $0 < |uw| < 0.04 \text{ m}^2 \text{ s}^{-2}$. This is despite the fact that they make a negative contribution to the stress. Correspondingly, there is little sediment movement associated with inward interactions ($u < 0$; $w < 0$). On average, sweeps and outward interactions contribute ~46% of the overall stress and 73% of the total acoustic intensity in only 24% of the time.

Table 1 Details of records used in comparisons of SGN acoustic intensity levels and contributions to the Reynolds stress (Fig. 3)

Record (h)	U_{100} (m s^{-1})	$-uw$ ($\text{m}^2 \text{ s}^{-2}$)	No. of events				Duration of events				Overall mean size (stress)				Overall mean size (SGN)				Contribution to stress			
			E	S	OI	II	E	S	OI	II	E	S	OI	II	E	S	OI	II	E	S	OI	II
0924-	1.23	0.0087	249	236	57	54	0.91	0.91	0.61	0.66	0.0201	0.0229	-0.0129	-0.0101	0.6345	2.2694	2.9384	0.5437	47.40	50.19	-4.21	-3.34
0944							(0.58)	(0.53)	(0.65)	(0.31)	(0.0104)	(0.0147)	(0.0058)	(0.0058)	(0.9329)	(3.1589)	(3.6808)	(0.6895)	(20.99)	(19.95)	(3.23)	(3.30)
0944-	1.27	0.0077	218	223	41	40	0.86	0.85	0.56	0.64	0.0188	0.0197	-0.0109	-0.0107	0.7936	1.4854	1.6443	0.6148	48.93	47.71	-2.96	-3.56
1001							(0.52)	(0.45)	(0.22)	(0.26)	(0.0109)	(0.0114)	(0.0045)	(0.0051)	(1.0129)	(1.4313)	(1.3794)	(0.7229)	(19.48)	(19.38)	(2.33)	(2.63)
1024-	1.16	0.0081	239	239	62	67	0.92	0.87	0.61	0.71	0.0207	0.0220	-0.0135	-0.0130	0.4452	1.3669	1.6727	0.5072	51.12	50.13	-5.22	-6.01
1044							(0.56)	(0.50)	(0.30)	(0.36)	(0.0104)	(0.0120)	(0.0047)	(0.0063)	(0.6413)	(1.8723)	(1.6221)	(1.2019)	(19.31)	(18.48)	(3.30)	(4.12)
1044-	1.11	0.0064	284	295	119	126	0.85	0.82	0.59	0.71	0.0145	0.0162	-0.0099	-0.0086	0.2776	0.6353	0.6661	0.2262	51.54	57.20	-8.69	-10.02
1104							(0.54)	(0.50)	(0.30)	(0.39)	(0.0101)	(0.0102)	(0.0070)	(0.0049)	(0.3392)	(0.8660)	(0.7863)	(0.3545)	(21.02)	(20.94)	(6.07)	(7.74)
1104-	1.08	0.0070	222	259	78	64	1.04	0.88	0.53	0.58	0.0163	0.0174	-0.0110	-0.0114	0.1129	0.4199	0.4098	0.1017	51.02	49.27	-5.34	-4.92
1124							(0.75)	(0.50)	(0.19)	(0.23)	(0.0082)	(0.0092)	(0.0060)	(0.0061)	(0.2220)	(0.6159)	(0.6224)	(0.1311)	(21.06)	(20.57)	(3.76)	(3.36)

U_{100} is the current at 1 m above the sea bed from near-bed velocity profile measurements. The water depth at this location was ~14 m. $-uw$ is the overall kinematic stress. The Reynolds stress is given by $\tau = -\rho uw$. All the records were obtained in the West Solent on 19 September 1982. Record lengths are 20 min with the exception of the record beginning 0944 h (BST) which is 17 min in length. E denotes ejections, S sweeps, OI outward interactions and II inward interactions. Values in parentheses are standard deviations except in the case of contributions to the stress where they indicate the proportions of time as a percentage.

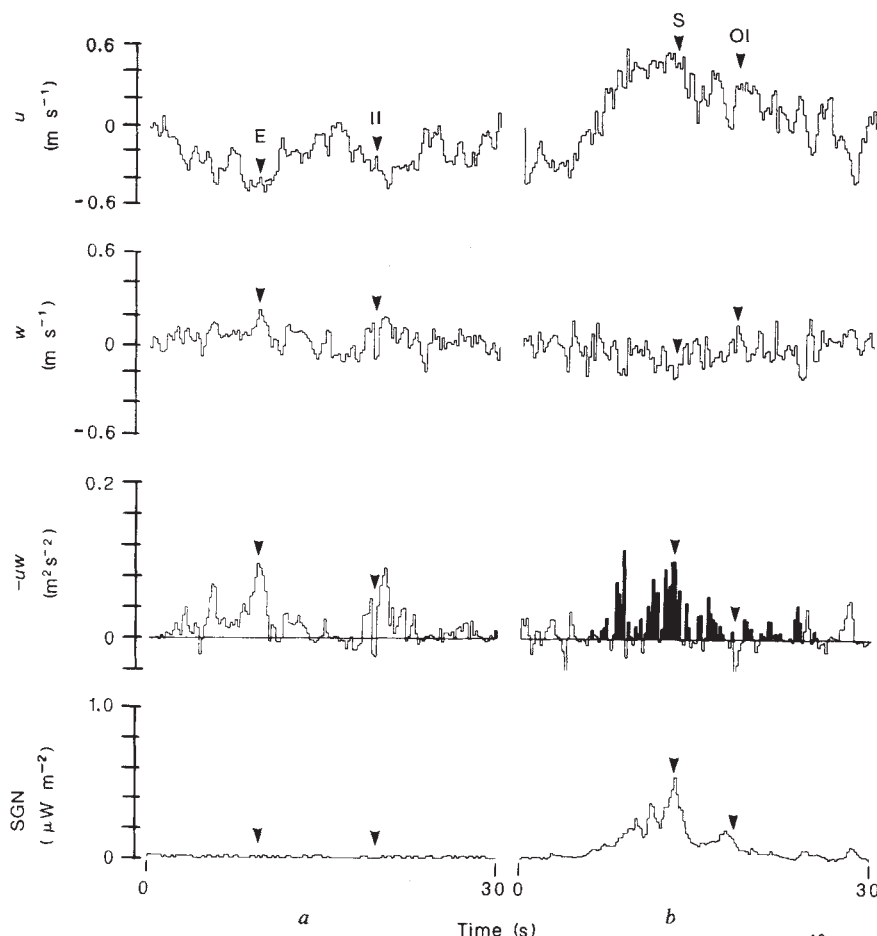
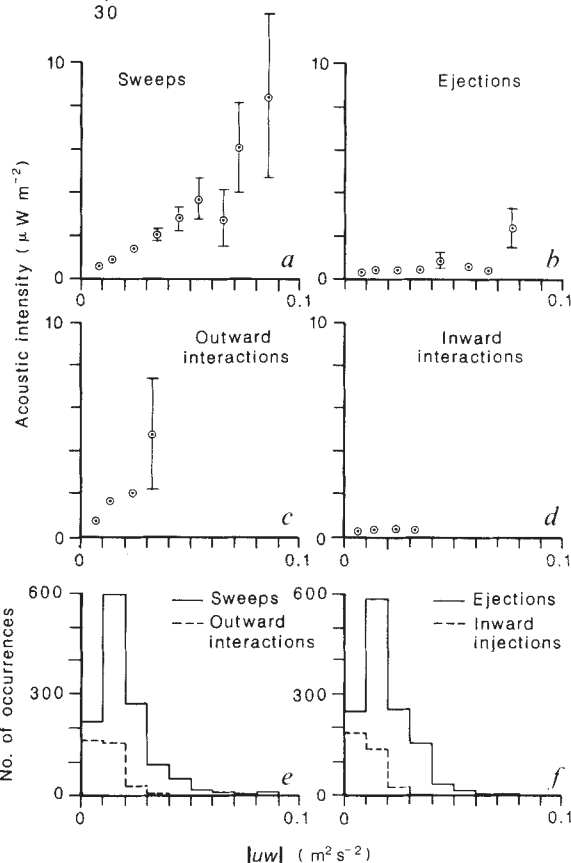


Fig. 2 (Left) u , w , $-uw$ and SGN records on an expanded timescale showing: *a*, ejection-type events (E) which produce no sediment movement; and *b*, similar-size sweep (S) events (black) producing appreciable sediment movement. Inward and outward interaction events, II and OI respectively, are also indicated. Only 5-Hz data are shown.

Fig. 3 (Right) Comparisons of sediment movement due to sweeps (*a*), ejections (*b*) and outward and inward interactions (*c* and *d*). *a*–*d* show mean values of the acoustic intensity in each $0.01 \text{ m}^2 \text{ s}^{-2}$ interval of $|uw|$, plotted against the mean stress in the same range. The number of events comprising each mean value of the acoustic intensity is indicated in *e* and *f*. Observations with <2 events are not shown. Error bars correspond to ± 1 s.e.m. and are not shown for standard errors $<0.2 \mu\text{W m}^{-2}$. Standard error bars for the stress were $<4\%$ of each mean and for clarity are not shown. The total number of events was 3,187.



These results suggest that horizontal turbulent velocity fluctuations (u) may have greater dynamical significance in terms of sediment movement, than the instantaneous contributions ($-uw$) to the Reynolds stress. However, vertical fluctuations (w) may still be important provided they are associated with increases in u . This seems to be the case for outward interactions (Fig. 3c), where $w > 0$ may indicate additional lift on exposed gravel particles by fluid moving away from the bed. This, in turn, may account for sediment transport rates which are higher than those in similar-size sweeps.

The observed close dependence on u could be explained if the gravel was moved principally by form drag acting on the flow-normal projected area of exposed particles. Sediment sampling²¹ has confirmed that the West Solent gravels are poorly sorted and underwater TV observations⁴ have shown that sediment movement is by sliding or rolling of exposed particles. To test this hypothesis, we have carried out correlation analyses on the acoustic intensity and turbulent flow records in Table 1, first in terms of the form drag proportional to $(\bar{U} + u)^2$, where $\bar{U}$ is the horizontal mean flow, and second in terms of contributions $|uw|$ to the Reynolds stress ($-\rho \overline{uw}$). The results of these comparisons are shown in Table 2 and indicate that in all cases sediment movement is better correlated with form drag than with the instantaneous stress. Note that although the correlation coefficients in Table 2 are low, they are, nevertheless, highly significant because of the very large number of data points in

each record (about 6,000). On average, Table 2 shows that the correlation (r) of sediment movement with $(\bar{U} + u)^2$ is a factor of two times better than that with $|uw|$.

Through the application of a novel acoustic technique, this work has provided a detailed high temporal resolution comparison of sediment response and turbulent flow conditions.

Table 2 Results of correlation analyses of SGN records with $(\bar{U}+u)^2$ and $|uw|$

Record (h)	Correlation coefficient (r) at 1 s lag SGN/ $(\bar{U}+u)^2$	SGN/ $ uw $
0924-0944	0.45	0.26
0944-1001	0.30	0.15
1024-1044	0.39	0.19
1044-1104	0.35	0.17
1104-1124	0.33	0.12

This has shown that bed-load movement of sea-bed gravels is caused principally by sweep-type motions in the bottom boundary layer and to a lesser extent by outward interactions. This observation can be explained if form drag rather than shear stress is assumed to be the principle cause of gravel movement. Furthermore, laboratory observations^{22,23} in steady currents have shown that the movement of finer sediments in suspension may also be linked to the turbulent bursting process. It is, therefore, interesting to speculate that a complete description of sediment movement in terms of the fluid turbulence might now be possible. In conceptual terms at least, this would consist of bed-load transport of coarse sediments driven principally by sweep-type events, while suspended load transport of finer sediments would be dominated by the ejection of low-momentum fluid away from the boundary. Beyond this, if sediment movement can be quantified in terms of the magnitude of sweep- and

ejection-type events, and if the frequency with which these motions occur can be predicted accurately⁹, then a new generation of sediment transport equations may be deduced which take the turbulent bursting process into account.

We thank J. M. Codd and A. A. Landa for assistance with computer programming and with the analysis of underwater TV data.

Received 11 March; accepted 10 June 1985.

1. Heathershaw, A. D. *Nature* **248**, 394-395 (1974).
2. Gordon, C. M. *Nature* **248**, 392-394 (1974).
3. Dyer, K. R. *Estuar. coast. mar. Sci.* **10**, 181-199 (1980).
4. Hammond, F. D. C., Heathershaw, A. D. & Longhorne, D. N. *Sedimentology* **31**, 51-62 (1984).
5. Corino, E. R. & Brodkey, R. S. *J. Fluid Mech.* **37**, 1-30 (1969).
6. Wallace, J. M., Eckelmann, H. & Brodkey, R. S. *J. Fluid Mech.* **54**, 39-48 (1972).
7. Kim, H. T., Kline, S. J. & Reynolds, W. C. *J. Fluid Mech.* **50**, 133-160 (1971).
8. Rao, K. N., Narasimha, R. & Narayanan, Badri, M. A. *J. Fluid Mech.* **48**, 339-352 (1971).
9. Blackwelder, R. F. & Haritonidis, J. H. *J. Fluid Mech.* **132**, 87-103 (1983).
10. Bradshaw, P. *Turbulence* (Springer, Berlin, 1976).
11. Jackson, R. G. *J. Fluid Mech.* **77**, 531-560 (1976).
12. Bridge, J. S. *Sedim. Geol.* **20**, 1-16 (1978).
13. Leeder, M. R. *Sedimentology* **30**, 485-491 (1983).
14. Allen, J. R. L. *Sedim. Geol.* **39**, 227-242 (1984).
15. Thorne, P. D., Heathershaw, A. D. & Troiano, L. *Mar. Geol.* **54**, M43-M48 (1984).
16. Thorne, P. D. *Mar. Geol.* (in the press).
17. Thorne, P. D., Salkield, A. P. & Marks, A. J. in *Acoustics and the Sea-Bed* (ed. Pace, N. C.) 395-402 (Bath University Press, 1983).
18. Heathershaw, A. D. *Geophys. J. R. astr. Soc.* **58**, 395-430 (1979).
19. Gordon, C. M. & Witting, J. in *Bottom Turbulence* (ed. Nihoul, J. C. J.) 59-81 (Elsevier, Amsterdam, 1977).
20. Soulsby, R. L. in *Physical Oceanography of Coastal and Shelf Seas* (ed. Johns, B.) 189-266 (Elsevier, Amsterdam, 1983).
21. Langhorne, D. N., Heathershaw, A. D. & Read, A. A. *Inst. oceanogr. Sci. Rep.* **140**, (1982).
22. Sumer, B. M. & Oguz, B. *J. Fluid. Mech.* **86**, 109-127 (1978).
23. Sumer, B. M. & Deigaard, R. *J. Fluid. Mech.* **109**, 311-337 (1981).

Hydrothermal particle plumes over the southern Juan de Fuca Ridge

Edward T. Baker, J. William Lavelle & Gary J. Massoth

Pacific Marine Environmental Laboratory, NOAA Building Number 3, 7600 Sand Point Way N.E., Seattle, Washington 98115, USA

Hydrothermal particles originate in buoyant plumes emanating from seafloor thermal vents¹⁻⁴ and accumulate as metalliferous sediments found along oceanic spreading centres⁵⁻⁷. Observational evidence of the transport pathway of hydrothermal particles is scarce, however, and the structure, extent and particle concentrations of hydrothermal plumes remain largely conjectural. To evaluate the potential of hydrothermal plumes as particle transport agents, we mapped their distribution along and across the 80-100-m-deep axial valley of the southern Juan de Fuca Ridge in the north-east Pacific Ocean (~44°30' N, 130° W), an area of known hydrothermal activity⁸⁻¹¹. Continuous light-scattering measurements, reported here, defined numerous particle plumes centred 30-120 m above bottom (m.a.b.); elemental analysis of the plume particles and the θ (potential temperature)- S (salinity) signatures of the plume water verified their hydrothermal origin. Concentrations of hydrothermal particles remained elevated at ridge-crest depths at least 100 km from the ridge, indicating distant off-axis dispersal of hydrothermal emissions.

Two-dimensional plume representations were constructed from continuous data acquired by an acoustically-tracked sensor package regularly cycled between 20 and 200 m.a.b. while being towed. Twenty-six tows along or across an 18-km segment of the axial valley were made between 27 May and 17 June 1984 (Fig. 1). Sensors on the tow body included an integrating scatterance meter (nephelometer), thermometer (resolution $\pm 0.0003^\circ\text{C}$), and conductivity cell (resolution $\pm 0.00004\text{ S m}^{-1}$). In addition to the tows, vertical conductivity-temperature-depth (CTD)/nephelometer/rosette casts over the ridge axis and up to 100 km off axis further defined plume distributions and collected particles for chemical analysis by secondary emission X-ray fluorescence¹² and for empirical calibration of the nephelometer. Relative light scattering values were linearly and significantly correlated (correlation coefficient, $r^2 = 0.86$) with mass con-

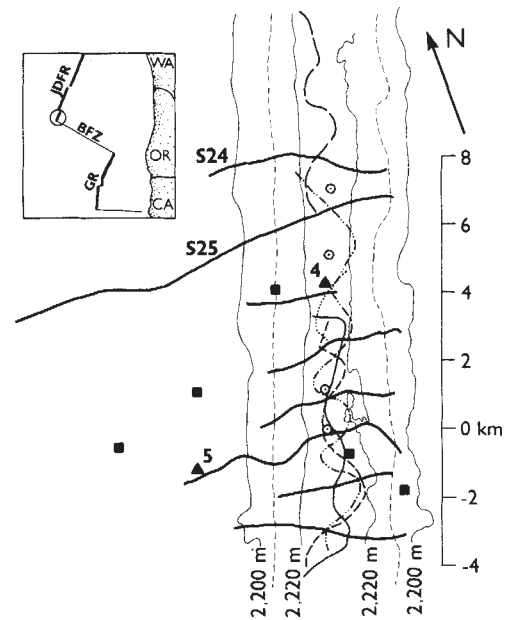


Fig. 1 The study area, circled in the inset, lies along the Gorda Ridge-Blanco Fracture Zone-Juan de Fuca Ridge rift system of the north-east Pacific Ocean. Selected tow tracks are shown relative to the floor of the flat axial valley (delimited by the 2,220-m contours) and the adjacent ridge crests (indicated by the thin dashed lines). The axial strike is $\sim 18^\circ$. Individual along-axis tow tracks can be identified by the key in Fig. 2b; cross-axis tracks are shown as thick solid lines. The locations of known active hydrothermal vents ($\odot$) as determined by deep-tow⁹ and submersible¹⁰ observations are shown along the centreline of the valley. Positions of moorings ($\blacktriangle$) and selected CTD casts ($\blacksquare$) are also shown. Sediment traps and current meters were located 50 m.a.b. on V4 and 100 m.a.b. on V5; an additional current meter was at 200 m.a.b. on V4.

centration determined by membrane filtration of 129 samples. Vertical particle flux (from sediment traps¹³) and current speeds were measured during the cruise on moorings in the axial valley (V4) and 4 km west of the valley (V5).

Each of the 26 tows found particulate plumes derived from hydrothermal emissions. The hydrothermal origin was indicated

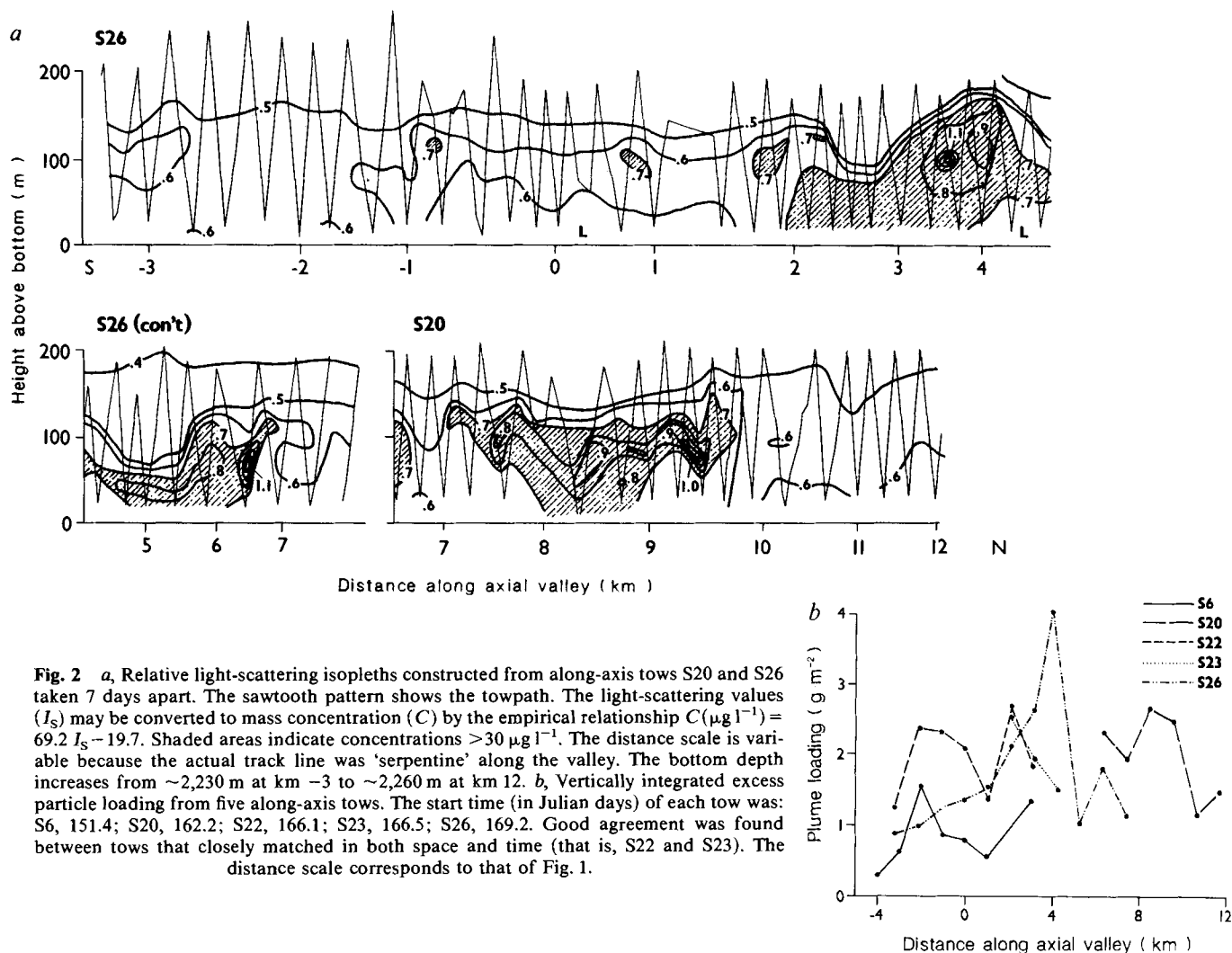


Fig. 2 *a*, Relative light-scattering isopleths constructed from along-axis tows S20 and S26 taken 7 days apart. The sawtooth pattern shows the towpath. The light-scattering values (I_s) may be converted to mass concentration (C) by the empirical relationship $C(\mu\text{g l}^{-1}) = 69.2 I_s - 19.7$. Shaded areas indicate concentrations $>30 \mu\text{g l}^{-1}$. The distance scale is variable because the actual track line was 'serpentine' along the valley. The bottom depth increases from $\sim 2,230$ m at km -3 to $\sim 2,260$ m at km 12. *b*, Vertically integrated excess particle loading from five along-axis tows. The start time (in Julian days) of each tow was: S6, 151.4; S20, 162.2; S22, 166.1; S23, 166.5; S26, 169.2. Good agreement was found between tows that closely matched in both space and time (that is, S22 and S23). The distance scale corresponds to that of Fig. 1.

by departures of up to 0.1°C from the local θ -S curve and (on casts) by particulate Fe/Al ratios enriched up to 79 times the background value of ~ 1 . Light-scattering observations, however, provided greater sensitivity than temperature in identifying hydrothermal plumes and far greater resolution than discrete chemical samples in describing them. Particle concentrations within the axial valley and off-axis plumes were elevated 2–10 times above the background particle concentration of $\sim 8 \mu\text{g l}^{-1}$ determined from mid-depth water samples. The horizon of maximum particle concentration in the axial valley characteristically occurred 30–120 m.a.b. (Fig. 2a) and was punctuated by localized maxima exceeding $35 \mu\text{g l}^{-1}$ and vertical gradients up to $5 \mu\text{g l}^{-1} \text{ m}^{-1}$. Well-defined plumelike features such as that centred at km 8.5 of tow S20 (Fig. 2a) were rare and restricted to a horizontal dimension of 1–3 km. Intense concentration maxima ($>40 \mu\text{g l}^{-1}$) preferentially occurred in the northern half of the study area and were separated by intervals of 2–3 km; this distribution conforms to the distribution of active vent fields documented by deep-tow photography⁹ and submersible observations¹⁰.

The concentration of particles originating solely from plume emissions was estimated by subtracting the background value of $8 \mu\text{g l}^{-1}$ from all transects and vertically integrating the resulting excess mass loading. These calculations will exceed the true emission loadings by the amount that resuspension of non-hydrothermal sediment contributes to the observed concentrations, but maximum current speeds of only 14 cm s^{-1} (V4) and the prevalence of above-bottom particle maxima suggest that resuspension was weak. Plume loadings calculated at 1-km intervals along five tows show marked spatial and temporal variability (Fig. 2b), consistent with the pattern expected from

discrete vent fields whose emissions are advected by a variable current regime. Temporal changes in discharge rates could also influence the interpretation of the observations.

After documenting the presence of hydrothermal plumes within the axial valley, we conducted eight cross-axis tows to evaluate the likelihood of off-axis dispersion of the particles. The tows were spaced at ~ 2 -km intervals along the valley (Fig. 1); all but the most southerly two found asymmetric plumes of varying intensity apparently originating along the centre or west side of the valley and extending up to, and sometimes over, the west ridge crest. (A particle plume at ridge-crest depths was also detected up to 10 km east of the valley by individual CTD casts.) Particle concentration isopleths along all of these tows roughly paralleled the bathymetry. The longest tow, S25, traced a continuous plume for >8 km west of the axial valley (Fig. 3a). The hydrothermal origin of this feature was confirmed by the continuity of elevated Fe/Al ratios of collected particles westward from the valley (Fig. 3a). Significant maxima at ridge-crest depths in both particle concentration and Fe/Al ratios were also observed on CTDs 20 km and 100 km to the west, suggesting that crestal emissions spread away from the ridge along the isopycnal $\sigma_\theta = 27.72$.

In contrast to these distinct off-axis trends, mean velocities at V4 and V5 during the cruise were weak ($1\text{--}2 \text{ cm s}^{-1}$) and directed northwards roughly parallel to the ridge axis. Mean cross-axis currents were $<5\%$ of along-axis currents within the valley (50 m.a.b. at V4) but increased to 25% 200 m.a.b. at V4 and to 20% at V5. Tidal currents (V4) were dominated by the M2 component, which had a major-axis amplitude of $\sim 4 \text{ cm s}^{-1}$, oriented along the ridge, with an ellipticity of ~ 0.1 .

The presence of a plume several kilometres west of the valley

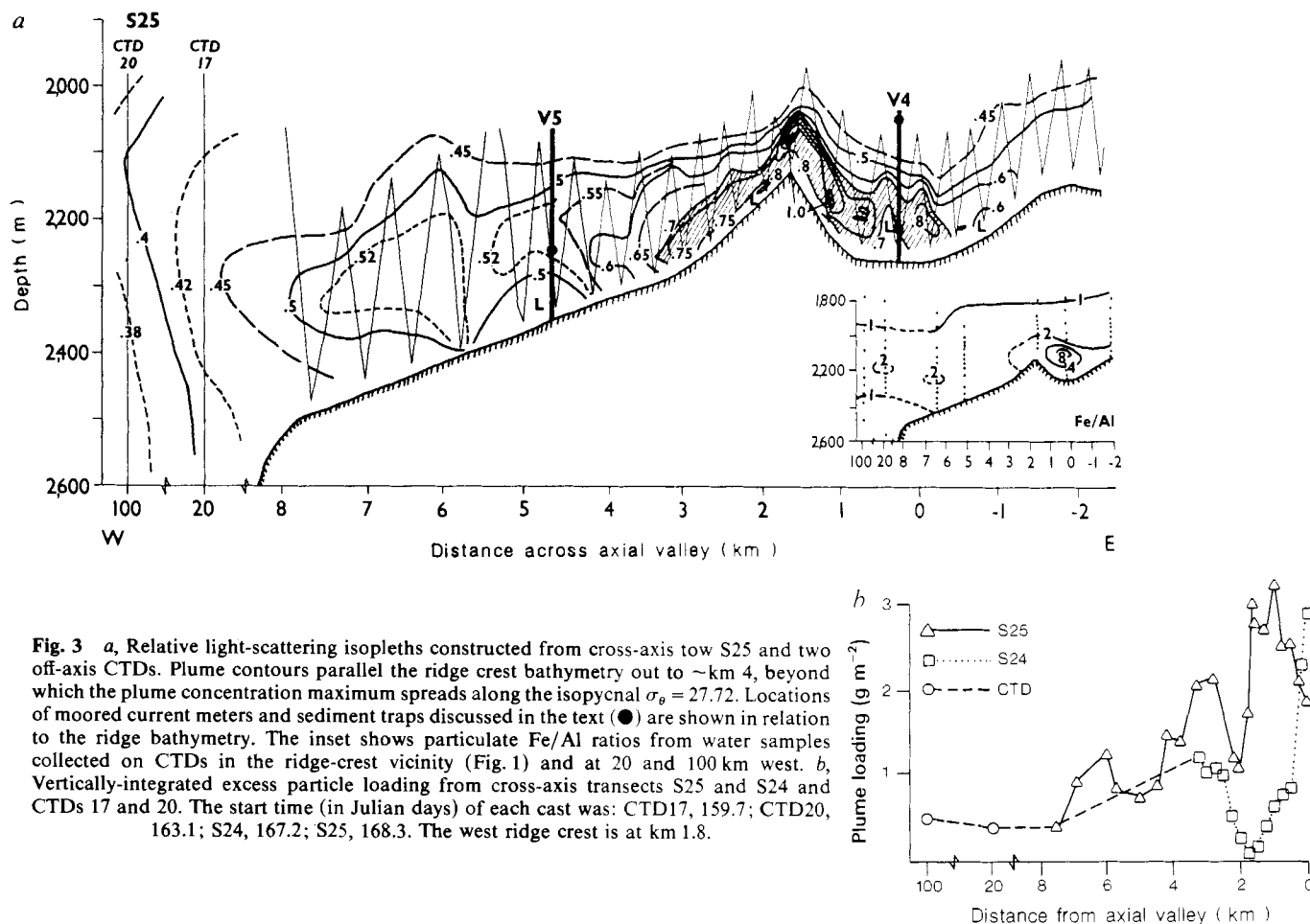


Fig. 3 *a*, Relative light-scattering isopleths constructed from cross-axis tow S25 and two off-axis CTDs. Plume contours parallel the ridge crest bathymetry out to ~km 4, beyond which the plume concentration maximum spreads along the isopycnal $\sigma_\theta = 27.72$. Locations of moored current meters and sediment traps discussed in the text (●) are shown in relation to the ridge bathymetry. The inset shows particulate Fe/Al ratios from water samples collected on CTDs in the ridge-crest vicinity (Fig. 1) and at 20 and 100 km west. *b*, Vertically-integrated excess particle loading from cross-axis transects S25 and S24 and CTDs 17 and 20. The start time (in Julian days) of each cast was: CTD17, 159.7; CTD20, 163.1; S24, 167.2; S25, 168.3. The west ridge crest is at km 1.8.

in the face of predominantly northward advection suggests that the plume had multiple sources. We speculate that particles from vent fields close to the S25 track caused the within-valley plume, while particles from more southerly vent fields crossed the track several kilometres west (and east) of the ridge crest. The vertically-integrated particle loading along S25 (Fig. 3*b*) shows a major peak on the western edge of the valley, and two lesser peaks west of the ridge crest that could be indicative of separate, southerly sources. Tow S24, taken 1 day earlier just to the north of S25, shows a similar loading increase, west of the ridge crest, that could have been continuous with the S25 plume (Fig. 3*b*). The paucity of our coverage, however, means that these results might have other interpretations.

If particle production by the vent fields was approximately steady state during the cruise, and if our off-axis transects were representative of the entire ridge, the off-axis loading decrease must result primarily from settling and/or horizontal dispersion. The contribution of settling can be evaluated from the vertical flux of hydrothermal material. The total particle flux into the sediment traps was $0.11 \pm 0.04 \text{ g m}^{-2} \text{ day}^{-1}$ for 10 two-day samples at V4, and $0.19 \text{ g m}^{-2} \text{ day}^{-1}$ for a single 10-day sample at V5. Elemental analysis indicates that the trapped material was predominantly biogenic and aluminosilicate, with <10% of the total attributable to hydrothermal phases. If a flux of $\leq 0.01 \text{ g m}^{-2} \text{ day}^{-1}$ is thus characteristic of hydrothermal phases in the study area, settling can have only a minor impact on the observed plume loadings. We note, however, that this flux could support the hydrothermal phase accumulation rates of $0.001\text{--}0.004 \text{ g m}^{-2} \text{ day}^{-1}$ found in the vicinity of the Juan de Fuca Ridge¹⁴.

We infer from these observations the following model of hydrothermal plume distributions around the southern Juan de Fuca Ridge. Buoyant vent emissions create a particle distribution that is vertically stratified around above-bottom concentration maxima. Along-axis variability arises because vent fields are not

continuous along the axial valley. Cross-axis tows and CTDs demonstrate that hydrothermal particles are effectively transported off axis, even though the net advective direction is northward, within roughly $\pm 20^\circ$ of the axial strike. We hypothesize that the loss rate of suspended hydrothermal phases by settling is slow enough to enable diffusion to coalesce individual plumes into a laterally uniform far-field hydrothermal cloud at distances of 20–100 km from the ridge crest, analogous to the cloud formed by dissolved tracers such as ^3He (ref. 15) and Mn (ref. 16).

This study has provided an initial description of the distribution of suspended hydrothermal particles; determination of a regional mass-balance between suspended and deposited particles will require quantification of both the percentage of emissions that escapes local deposition and becomes suspended in a plume, and the rate of *in situ* particle production by slow precipitation of dissolved species within the advecting plume.

We thank H. Milburn and R. Newman for technical assistance in the development of the towed sensor system. This research was supported by the NOAA VENTS Program.

Received 19 February; accepted 21 May 1985.

1. Corliss, J. B. *et al.* *Science* **203**, 1073–1083 (1979).
2. Lupton, J. E. *et al.* *Earth planet. Sci. Lett.* **50**, 115–127 (1980).
3. Edmond, J. M., Von Damm, K. L., McDuff, R. E. & Measures, C. I. *Nature* **297**, 187–191 (1982).
4. Campbell, A. C. & Gieskes, J. M. *Earth planet. Sci. Lett.* **68**, 57–72 (1984).
5. Bostrom, K. & Peterson, M. N. A. *Econ. Geol.* **61**, 1258–1265 (1966).
6. Bostrom, K., Peterson, M. N. A., Johnson, O. & Fisher, D. E. *J. geophys. Res.* **74**, 3261–3270 (1969).
7. Dymond, J. *Mem. geol. Soc. Am.* **154**, 133–173 (1981).
8. Delaney, J. R., Johnson, M. P. & Karsten, J. L. *J. geophys. Res.* **86**, 11747–11750 (1981).
9. Normark, W. R. *et al.* *Geology* **11**, 158–163 (1983).
10. Normark, W. R. *et al.* *EOS* **66**, 116 (1985).
11. Crane, K. *et al.* *J. geophys. Res.* **90**, 727–744 (1985).
12. Feely, R. A., Massoth, G. J. & Landing, W. M. *Mar. Chem.* **10**, 431–453 (1981).
13. Baker, E. T. & Milburn, H. B. *Cont. Shelf Res.* **1**, 425–435 (1983).
14. Leinen, M. *EOS* **65**, 1112 (1984).
15. Lupton, J. E. & Craig, H. *Science* **214**, 13–18 (1981).
16. Klinkhammer, G. P. *Chem. Geol.* **29**, 211–226 (1980).

Acetylcholine activates an inward current in single mammalian smooth muscle cells

C. D. Benham, T. B. Bolton & R. J. Lang

Department of Pharmacology and Clinical Pharmacology,
St George's Hospital Medical School, London SW17 0RE, UK

Acetylcholine, the major excitatory neurotransmitter to the smooth muscle of mammalian intestine¹, is known to depolarize smooth muscle cells with an apparent increase in membrane conductance². However, the ionic mechanisms that are triggered by muscarinic receptor activation and underlie this response are poorly understood, due in part to the technical problems associated with the electrophysiological study of smooth muscle³. The muscarinic action of acetylcholine in certain neurones has been shown to involve the switching off of a resting K^+ current (M-current)⁴ and a similar mechanism has recently also been identified in smooth muscle of amphibian stomach⁵. We have now applied the patch-clamp technique⁶ to single smooth muscle cells^{7,8} of rabbit jejunum and find that muscarinic receptor activation switches on a non-selective, voltage-sensitive inward current. In addition, acetylcholine activates and then suppresses spontaneous K^+ current transients, which are probably triggered by rises in intracellular Ca^{2+} in these cells.

Membrane current and voltage recordings were made from enzymatically dispersed single cells isolated from the longitudinal smooth muscle layer of adult rabbit jejunum⁷. These cells were 5–15 μm wide, up to 300 μm long and contracted on application of acetylcholine. Acetylcholine was applied either in the bathing solution or by iontophoresis from a glass micro-pipette placed about 10 μm from the mid-region of partially contracted cells, chosen to ensure uniform potential during voltage clamp (Fig. 1). Membrane currents and membrane potentials were recorded with a glass pipette in the 'whole-cell' recording configuration⁶.

In normal physiological saline solution, acetylcholine applied iontophoretically or in the bathing solution (0.1–10 μM) evoked contraction, membrane depolarization associated with an increase in membrane conductance (indicated by a reduction in size of hyperpolarizing electrotonic potentials) and action potential discharge (Fig. 1a). Thus, these single cells show similar membrane potential responses to those of cells within intact strips of smooth muscle tissue recorded with other techniques². During voltage-clamp, acetylcholine evoked an inward current at holding potentials between –40 and –80 mV (Figs 1b, 2a,b). This inward current cannot be explained as a decrease of a resting K^+ current, as input resistance decreases and there is an increase in current noise indicating that additional channels were being opened rather than already open ones being closed (Fig. 1a, c). To further investigate the mechanism of action of acetylcholine, we applied it to cells in solutions with elevated K^+ concentrations clamped at the K^+ equilibrium potential, E_K , so eliminating any possibility of a contribution from changes in K currents under these conditions. Figure 1c shows a response obtained from a cell bathed in 45 mM K^+ and voltage clamped at E_K (–25 mV). A large inward current was still generated by acetylcholine application, showing that inward current was activated which at this potential could not be due to K^+ ion flow.

Further analysis of the ionic selectivity of this current required the measurement of responses from the same cell at various membrane potentials, and this was made more difficult by the slowness of the response and its decline in size upon repetition. However, by bathing cells in elevated K^+ -containing solutions, we were able to obtain stable recordings for up to 40 min that allowed several responses such as those shown in Fig. 2a to be recorded in the same cell. At E_K (–25 mV) the inward current response was close to its maximum of about 30 pA after about

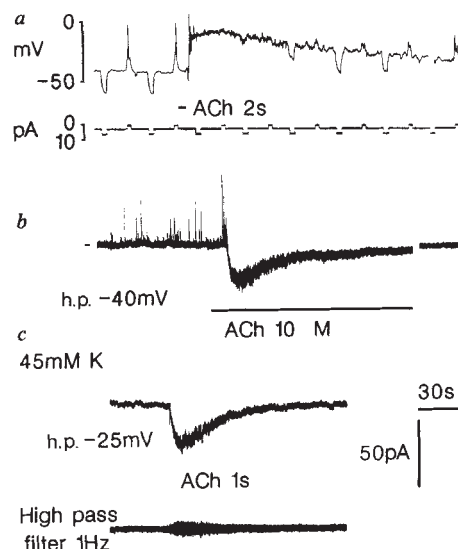


Fig. 1 Membrane potential and membrane current responses of jejunal smooth muscle cells to acetylcholine (ACh) application. **a**, Membrane potential record from a single cell with a small holding current applied to suppress spontaneous action potential discharge. Current pulses of alternating polarity were injected to evoke electrotonic potentials, and the depolarizing pulses elicited action potentials. ACh (2 s, 50 nA) was applied from an iontophoretic pipette positioned about 10 μm from the mid-region of the cell. This cell had an input resistance of 4.5 G Ω and cell capacitance of 70 pF (calculated from the hyperpolarizing electrotonic potential), which indicates a surface area of 6,700 μm^2 if 10^{-6} F cm^{-2} is assumed for specific membrane capacitance. This gives a space constant at the resting membrane resistance of about 5 mm, which is 60 times greater than the length of the short cable in these single cells. Thus, cells under voltage-clamp will be close to isopotential even with a substantial increase in conductance during acetylcholine application. The cell was bathed in a normal physiological saline solution containing (mM) Na^+ 125; K^+ 6; Ca^{2+} 2.5; Mg^{2+} 1.2; Cl^- 133; HPO_4^- 1.2; HEPES 10; glucose 11. The pipette (intracellular) solution contained (mM) Na^+ 5; K^+ 126; Mg^{2+} 1.2; Cl^- 128; HPO_4^- 1.2; EGTA 0.77; glucose 11; HEPES 10. All solutions were buffered to pH 7.2 and the temperature was 25 $^{\circ}C$. **b**, Membrane current response of a cell under voltage-clamp to bath application of acetylcholine when the holding potential (h.p.) was –40 mV. Cell was bathed in same solution as **a**. **c**, Membrane current response of another cell bathed in 45 mM K^+ solution (K^+ replacing Na^+) at the calculated K^+ equilibrium potential (–25 mV). The lower trace shows the same response high-pass filtered at 1 Hz to show the increase in current noise on acetylcholine application. This increase in current noise is most simply explained by the opening and closing of channels activated by acetylcholine responsible for the inward current.

5 s and had a latency of around 500 ms. This current was reduced in amplitude at both more negative and more positive potentials. Figure 2b shows the current needed to clamp the cell at various potentials before and during the effects of acetylcholine application. The two extrapolated lines intersect just positive to 0 mV. This extrapolated E_{rev} does not correspond to the equilibrium potential of any one ion, indicating that the current is carried by a mixture of ions—probably K^+ , Na^+ and/or Ca^{2+} as has been concluded from work on whole smooth muscle tissue^{1,2}. The net agonist-induced current plotted against potential in this elevated K^+ concentration was U-shaped due to deviation from an ohmic conductance at potentials negative to –20 mV (Fig. 2c). Also shown are the averaged responses of 15 cells in normal (6 mM K^+) solution. The shape of the relationship appeared to be similar although shifted to the left, as might be expected from a leftward shift in E_K if K^+ was a permeant ion. Records obtained in this solution in the most interesting range of –30 to 0 mV, where the reversal potential probably lies, were complicated by the appearance of a burst of transient outward currents upon acetylcholine application (see below).

In solutions containing 45 mM K^+ , we had a better estimate

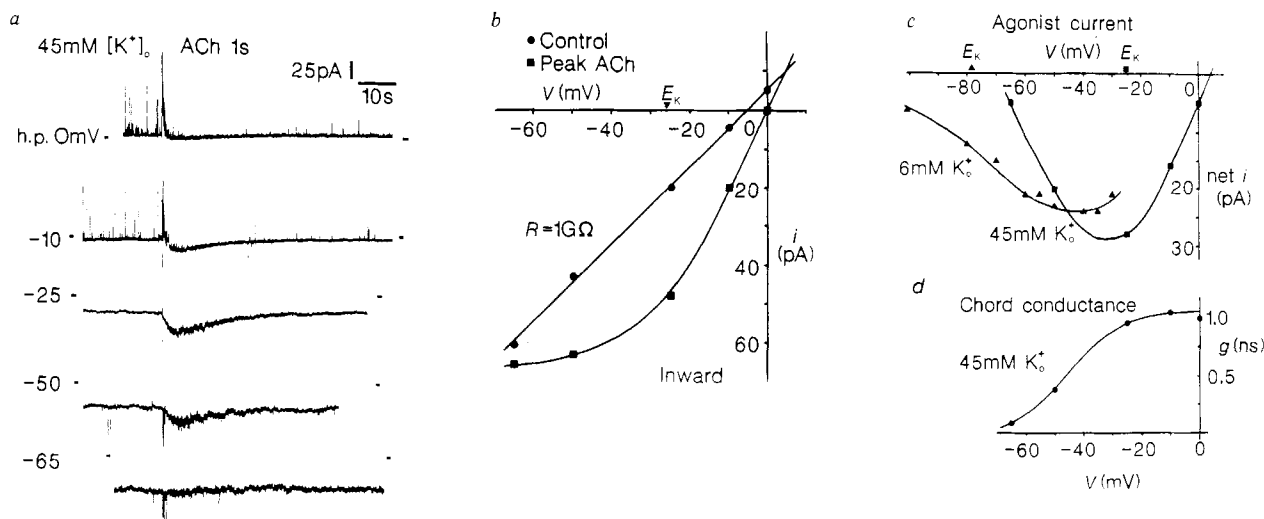


Fig. 2 *a*, Membrane current responses recorded from a single cell under voltage-clamp at five different holding potentials. Responses to 1-s iontophoretic pulses of acetylcholine were evoked at intervals of 8 min. Small horizontal bars indicate zero current level. Note the transient currents that appear as vertical deflections at holding potentials (h.p.) of 0, -10 mV (when they are outward) and -50, -65 mV (when they are inward); they are not seen at E_K (-25 mV). On acetylcholine application their frequency is temporarily increased and they may summate to produce an initial K^+ current that precedes the slower acetylcholine-induced current; after their initial potentiation the transient currents are inhibited for minutes. *b*, Current required to hold this cell at various potentials in the absence and presence of acetylcholine plotted from responses shown in *a*. The acetylcholine-induced current is maximal close to E_K and declines at potentials positive and negative to E_K ; it has an extrapolated reversal potential of about +5 mV. *c*, Net agonist-induced current at various potentials plotted from responses shown in *a* (■), and means of 33 responses from 15 cells in normal physiological saline (▲) $n = 2-7$. *d*, Chord conductance of responses shown in *a* using the extrapolated reversal potential under these conditions of +5 mV. The slope of the conductance curve between -50 and -25 mV gave an e-fold shift per -27 mV.

of E_{rev} at +5 mV (Fig. 2c) and this enabled the chord conductance of the current to be calculated. This showed that the acetylcholine-evoked current increased over the range -70 to -20 mV. The same type of voltage sensitivity, which could produce a regenerative depolarization, is produced by glutamate acting on *N*-methyl-D-aspartate (NMDA) receptors in certain neurones⁹. The non-ohmic properties of the NMDA current have recently been shown to be due to block of the channels at negative potentials by Mg^{2+} ions^{10,11}. We have not observed any effect of removing external Mg^{2+} on responses to acetylcholine, which suggests that the voltage dependence was not due to Mg^{2+} block. When instantaneous hyperpolarizing voltage jumps from -40 to -70 mV were applied at the peak of the acetylcholine response, an inward current was observed that decayed to a steady state with a $t_{1/2}$ of about 5 ms.

Acetylcholine application produced other changes in membrane currents. Spontaneous transient outward currents of varying amplitude, about 50 ms duration and with a frequency of 0.5-3 Hz were observed in cells bathed in elevated K^+ solution with normal calcium concentration and also in cells in normal solution at depolarized potentials, conditions which would tend to cause an increased level of internal calcium. Application of acetylcholine sometimes induced a brief potentiation of these transient outward currents in both frequency and amplitude. This potentiation preceded the onset of the inward current, and was consistently followed by long-lasting inhibition of the transient outward currents (Fig. 2a). These brief currents may represent the activation in response to transient increases in intracellular calcium of Ca^{2+} -activated K^+ channels which we have identified in these cells¹²; the transient currents reversed at the K^+ equilibrium potential (Fig. 2a) were blocked by external Ba^{2+} ions (2-10 mM) and were not seen if Cs^+ replaced K^+ in the pipette or if the concentration of EGTA in the pipette was raised to 10 mM. These currents were activated by caffeine (5 mM), which is known to release intracellular stores of Ca^{2+} in muscle¹³. Fluctuations in intracellular calcium concentration ($[Ca^{2+}]_i$) that might stimulate such currents have been measured directly in cardiac cells¹⁴, and similar transient potassium currents caused by fluctuations in $[Ca^{2+}]_i$ have been observed in neurones¹⁵.

A biochemical mediator within the cell may be important in triggering the inward current produced by muscarinic receptor activation, as the long latency of this response suggests a rather indirect coupling between receptor and channels. The latency can no longer be dismissed as being due to a long extracellular diffusion pathway, as in our system we know the cells must have clean, naked membranes to obtain gigaohm seals with our recording pipette. (A diffusion distance of 10 μ m would induce a nicotinic response at the neuromuscular junction in less than 10 ms¹⁶. A rise in $[Ca^{2+}]_i$ does not seem to be a possible mechanism (as in neurones¹⁷ and the heart¹⁸), because caffeine did not stimulate an inward current.

In conclusion, we have succeeded in dissecting some of the complex membrane events that underlie the excitatory muscarinic action of acetylcholine in mammalian smooth muscle. The major effect of acetylcholine was to stimulate a voltage-dependent inward current that leads to depolarization of the cell and initiates action potential discharge. This current was large at the potassium equilibrium potential and was a non-selective increase in permeability, probably to cations, as has been described for whole smooth muscle tissues^{1,2,19}. In addition, acetylcholine causes a long-lasting inhibition of spontaneous outward current transients, which are probably caused by the transient opening of Ca^{2+} -activated K^+ channels. The two types of channel modulation that we have observed are sufficient to explain the membrane responses of whole tissues of mammalian visceral smooth muscle to acetylcholine.

This work was supported by the MRC.

Received 17 January; accepted 29 May 1985.

- Bolton, T. B. *Br. med. Bull.* **5**, 275-283 (1979).
- Bolton, T. B. *J. Physiol., Lond.* **220**, 647-671 (1972).
- Bolton, T. B., Tomita, T. & Vassort, G. in *Smooth Muscle: An Assessment of Current Knowledge* (eds Bülbring, E. et al.) (Edward Arnold, London).
- Brown, D. A. & Adams, P. R. *Nature* **283**, 673-676 (1980).
- Sims, S. M., Singer, J. J. & Walsh, J. V. *Soc. Neurosci. Abstr.* **9**, 732 (1983).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
- Benham, C. D. & Bolton, T. B. *J. Physiol., Lond.* **340**, 469-486 (1983).
- Benham, C. D., Bolton, T. B. & Lang, R. J. *J. Physiol., Lond.* **353**, 67P (1984).
- Mayer, M. L. & Westbrook, G. J. *J. Physiol., Lond.* **354**, 29-54 (1984).
- Nowak, L., Bregeotovski, P., Ascher, P., Herbert, A. & Prochiantz, A. *Nature* **307**, 462-465 (1984).

11. Mayer, M. L., Westbrook, G. & Guthrie, P. B. *Nature* **309**, 261–263 (1984).
12. Benham, C. D., Bolton, T. B., Lang, R. J. & Takewaki, T. *J. Physiol., Lond.* (in the press).
13. Fabiato, A. & Fabiato, F. *Circulation Res.* **40**, 119–129 (1977).
14. Orchard, C. H., Eisner, D. A. & Allen, D. G. *Nature* **304**, 735–738 (1983).
15. Brown, D. A., Constanti, A. & Adams, P. R. *Cell Calcium* **4**, 407–420 (1983).
16. Katz, B. & Miledi, R. *Proc. R. Soc. Lond. B* **161**, 483–495 (1965).
17. Yellen, G. *Nature* **296**, 357–359 (1982).
18. Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
19. Bolton, T. B. *Physiol. Rev.* **59**, 606–718 (1979).

Inositol 1,4,5-trisphosphate induces calcium release from sarcoplasmic reticulum of skeletal muscle

Pompeo Volpe^{*†}, Giovanni Salvati^{*†},
Francesco Di Virgilio^{†‡} & Tullio Pozzan[†]

^{*} Centro di Studio per la Biologia e la Fisiopatologia Muscolare del CNR, [†] Centro di Studio per la Fisiologia dei Mitochondri del CNR and [‡] Istituto di Patologia Generale dell'Università di Padova, via Loredan 16, 35131 Padova, Italy

The sarcoplasmic reticulum of skeletal muscle is a specialized form of endoplasmic reticulum¹ that controls myoplasmic calcium concentration and, therefore, the contraction–relaxation cycle². Ultrastructural studies³ have shown that the sarcoplasmic reticulum is a continuous but heterogeneous membranous network composed of longitudinal tubules that surround myofibrils and terminal cisternae. These cisternae are junctionally associated, via bridging structures called 'feet'⁴, with sarcolemmal invaginations (the transverse tubules) to form the triadic junction⁴. Following transverse tubule depolarization, a signal, transmitted along the triadic junction, triggers Ca^{2+} release from terminal cisternae^{5,6}, but the mechanism of this coupling is still unknown⁷. Inositol 1,4,5-trisphosphate ($\text{Ins}(1,4,5)\text{P}_3$) has recently been shown to mobilize Ca^{2+} from intracellular stores, referable to endoplasmic reticulum, in a variety of cell types (see ref. 8 for review), including smooth muscle cells of the porcine coronary artery⁹ and canine cardiac muscle cells¹⁰. Here we show that $\text{Ins}(1,4,5)\text{P}_3$ (1) releases Ca^{2+} from isolated, purified sarcoplasmic reticulum fractions of rabbit fast-twitch skeletal muscle, the effect being more pronounced on a fraction of terminal cisternae that contains morphologically intact feet structures¹¹; and (2) elicits isometric force development in chemically skinned muscle fibres.

Ca^{2+} uptake by and Ca^{2+} release from isolated sarcoplasmic reticulum (SR) fractions were measured with a Ca^{2+} -sensitive electrode. Figure 1a shows that terminal cisternae (TC) accumulated 43 nmol of Ca^{2+} per mg of protein and then (dashed line) reached a slightly lower steady state¹². Addition of 25 μM $\text{Ins}(1,4,5)\text{P}_3$ (continuous line) promoted Ca^{2+} release, and a new, much lower steady state was attained. Following a biphasic pattern, approximately 50% of the accumulated Ca^{2+} was released within 5 min and then slowly re-accumulated. The effect of $\text{Ins}(1,4,5)\text{P}_3$ was specific since neither inositol 1,4-bisphosphate ($\text{Ins}(1,4)\text{P}_2$) nor inositol 4,5-bisphosphate ($\text{Ins}(4,5)\text{P}_2$) caused Ca^{2+} release from either TC fractions or fractions of longitudinal tubules (LSR), when tested at 25 μM (Table 1).

Figure 1b shows the relationship between the extent of Ca^{2+} release and $\text{Ins}(1,4,5)\text{P}_3$ concentration. Half-maximal stimulation of Ca^{2+} release was obtained at 4–5 μM $\text{Ins}(1,4,5)\text{P}_3$ for both TC and LSR fractions. $\text{Ins}(1,4,5)\text{P}_3$ was more effective on TC than on LSR, where only 20% of the accumulated Ca^{2+} was released. The partial sensitivity of LSR to $\text{Ins}(1,4,5)\text{P}_3$ is due, at least in part, to the contamination of this fraction by about 20% TC vesicles (compare with ref. 11), as indicated by the effect of ruthenium red, an inhibitor of Ca^{2+} release from TC^{13–15} but not from LSR^{15,16} fractions. Ruthenium red at 20 μM almost completely blocked $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release from TC, but produced a decrease of only 50% in release from LSR

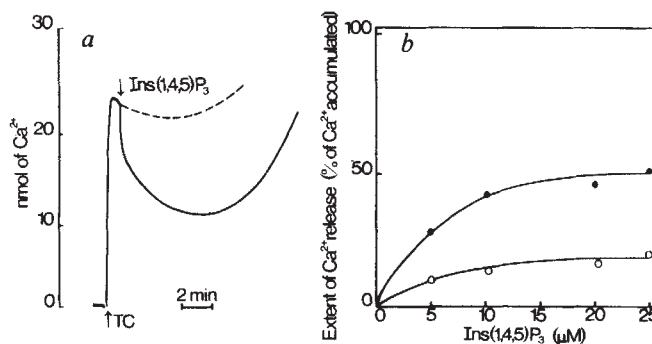


Fig. 1 $\text{Ins}(1,4,5)\text{P}_3$ promotes Ca^{2+} release from isolated SR fractions. *a*, After peak Ca^{2+} accumulation, there was a small, basal Ca^{2+} release (dashed line). Addition of 25 μM $\text{Ins}(1,4,5)\text{P}_3$ (continuous line) promoted a large Ca^{2+} release from TC, and a new lower steady state was attained. The rate of Ca^{2+} release induced by $\text{Ins}(1,4,5)\text{P}_3$ (fast phase) could not be determined accurately because of the inherent slow response of the Ca^{2+} electrode. As a lower limit, we calculated a rate of 1 nmol of Ca^{2+} per s per mg of TC protein, that is, three orders of magnitude lower than that estimated *in vivo*³¹. $\text{Ins}(1,4,5)\text{P}_3$ by itself produced only a small increase of medium Ca^{2+} that accounted for about 4% of the effect shown. *b*, A summary of data obtained from several experiments as shown in *a*, using either LSR (○) or TC (●) fractions. The extent of $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release was calculated as the difference between total (+ $\text{Ins}(1,4,5)\text{P}_3$) and basal Ca^{2+} release at steady state, and was expressed as percentage of Ca^{2+} accumulated by the SR. Similar results were obtained using three different SR preparations.

Methods. SR fractions were isolated from rabbit fast-twitch skeletal muscles and purified by differential and density gradient centrifugation^{11,32}. R2 (LSR) and R4 (TC) fractions were resuspended in 0.3 M sucrose, 5 mM imidazole pH 7.4 and stored at -70°C until used. Protein concentration was determined by the Lowry method³³, using bovine serum albumin as a standard. $\text{Ins}(1,4,5)\text{P}_3$ was prepared by incubating human erythrocyte ghosts at 37°C for 15 min with 2 mM CaCl_2 followed by Dowex-formate column separation, salt removal with 2 M LiCl and three consecutive extractions with ethanol^{24,34}. Ca^{2+} uptake and Ca^{2+} release were continuously monitored with a Ca^{2+} -sensitive electrode (G. Moller, Zurich). The assay was carried out at room temperature ($22\text{--}24^\circ\text{C}$) in a medium containing, in a final volume of 4 ml, 20 mM Tris-Cl pH 6.8, 0.1 M KCl, 5 mM MgSO_4 , 2.5 mM ATP, 5 mM $\text{Na}_2\text{-phosphocreatine}$ and 80 μg creatine kinase. Several pulses of 10 nmol CaCl_2 were sequentially added to calibrate the response of the Ca^{2+} electrode, and the reaction was started with either 600 μg of TC or 300 μg of LSR protein, when medium-free Ca^{2+} was 4 μM LSR and TC fractions accumulated 102 ± 9 and 40 ± 5 nmol of Ca^{2+} per mg of protein, respectively (mean $\pm$ s.d.; $n = 5$).

(Table 1). The small ruthenium red-insensitive $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release appears to be a LSR characteristic (Table 1). Thus, our results suggest that $\text{Ins}(1,4,5)\text{P}_3$ acts preferentially on junctional SR (TC) fractions. Interestingly, Berridge and Irvine⁸ have postulated a subspecialization for endoplasmic reticulum of other cell types in which only regions lying close to the plasma membrane might be sensitive to $\text{Ins}(1,4,5)\text{P}_3$.

The ability of $\text{Ins}(1,4,5)\text{P}_3$ to release Ca^{2+} from the SR was further investigated using a chemically skinned fibre preparation¹⁷. Ca^{2+} release from the SR was measured indirectly by monitoring isometric force development¹⁷. Segments of single fibres (average mean diameter 50 μm) were exposed to Ca^{2+} loading solution (free $\text{Ca}^{2+} = 0.40 \mu\text{M}$) for 30 s, so that the SR could accumulate Ca^{2+} , rinsed in EGTA-free solution and then challenged with $\text{Ins}(1,4,5)\text{P}_3$ that elicited transient force development (Fig. 2a). At low $\text{Ins}(1,4,5)\text{P}_3$ concentrations (1 μM , upper trace in Fig. 2a), the time course of force development was biphasic, that is, a slow phase turned into a faster one after approximately 30 s. An S-shaped time course suggests that $\Delta[\text{Ca}^{2+}]/\Delta t$ in the intermyofibrillar space was very small¹⁸ after stimulation. At higher $\text{Ins}(1,4,5)\text{P}_3$ concentrations, (20 μM , lower trace in Fig. 2a), force development was monophasic. These results imply that $\text{Ins}(1,4,5)\text{P}_3$ was able to release Ca^{2+}

from TC, which are the only Ca^{2+} storage compartments of the SR of skeletal muscle fibres^{6,19}. At $20 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$, $t_{1/2}$ to peak tension was 3 s and rate of tension rise, normalized against that obtained in the presence of 10 mM caffeine, a drug known to cause Ca^{2+} release from TC of isolated fractions²⁰ and of intact²¹ and skinned²² fibres, was 0.58. The slow rate of force development might be determined by the rate of $\text{Ins}(1,4,5)\text{P}_3$ diffusion throughout the cross-section of the fibre, and by the extent of $\text{Ins}(1,4,5)\text{P}_3$ hydrolysis (see below). Figure 2b shows a plot of relative rate of tension rise versus $\text{Ins}(1,4,5)\text{P}_3$ concentration. Half-maximal rate was obtained at $3 \mu\text{M}$. $\text{Ins}(1,4,5)\text{P}_3$ action depended critically on Ca^{2+} loading by the SR of skinned fibres. The effect of $10 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$ on the rate of tension rise was decreased by 70% when loading time was reduced from 30 to 15 s or was abolished if the free Ca^{2+} of the loading solution was decreased from 0.4 to $0.16 \mu\text{M}$.

As with isolated SR fractions, the ability to release Ca^{2+} appears to be specific for $\text{Ins}(1,4,5)\text{P}_3$, since $20 \mu\text{M}$ of either $\text{Ins}(1,4)\text{P}_2$ or $\text{Ins}(4,5)\text{P}_2$ had a minimal effect on force development in skinned fibres, that is 3–6% of that obtained with the same amount of $\text{Ins}(1,4,5)\text{P}_3$ (not shown). Furthermore, $\text{Ins}(1,4,5)\text{P}_3$ -induced force development was completely abolished by $20 \mu\text{M}$ ruthenium red (not shown). Ruthenium red inhibits Ca^{2+} release from the SR of chemically skinned fibres (G.S. and P.V., in preparation) as well as from isolated TC fractions.

The proposed role of $\text{Ins}(1,4,5)\text{P}_3$ as a second messenger for releasing Ca^{2+} from intracellular stores⁸ requires: first, substrates (phosphatidylinositol 4,5-bisphosphate, $\text{PtdIns}(4,5)\text{P}_2$), specific hydrolases ($\text{PtdIns}(4,5)\text{P}_2$ phosphodiesterase or phospholipase C) at the surface membrane level, and inactivating enzymes (inositol triphosphatase); second, specificity for endoplasmic reticulum receptors; and third, a polyphosphoinositide breakdown rate compatible with cellular response times.

The data reported here are consistent with the possibility that $\text{Ins}(1,4,5)\text{P}_3$ is the specific chemical transmitter in the triadic junction where signal transduction for muscle activation occurs. In fact: (1) $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release was preferentially displayed by TC fractions, which are enriched in junctional SR membranes, and was antagonized by ruthenium red, a blocker of TC Ca^{2+} channels¹³; (2) $\text{Ins}(1,4,5)\text{P}_3$ evoked Ca^{2+} release from TC of skinned muscle fibres; and (3) $\text{Ins}(1,4,5)\text{P}_3$ was preferred over $\text{Ins}(1,4)\text{P}_2$ and $\text{Ins}(4,5)\text{P}_2$. A simplified model for transverse tubule (TT)–TC coupling may be outlined as follows. After TT depolarization, $\text{Ins}(1,4,5)\text{P}_3$ is produced at the myoplasmic leaflet of the TT membrane and released into the triadic junction; $\text{Ins}(1,4,5)\text{P}_3$ -sensitive Ca^{2+} channels, localized in junctional SR (TC) membranes, open and myoplasmic free Ca^{2+} rises. Alternatively, $\text{Ins}(1,4,5)\text{P}_3$ may evoke a small Ca^{2+}

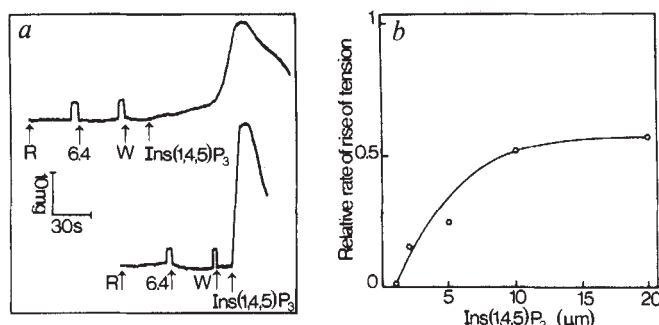


Fig. 2 $\text{Ins}(1,4,5)\text{P}_3$ elicits isometric force development in chemically skinned muscle fibres. **a**, Upper and lower traces show the effect of 1 and $20 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$, respectively. **b**, A summary of data from several experiments as shown in **a**. Rate of tension rise, at specified $\text{Ins}(1,4,5)\text{P}_3$ concentration, was normalized against that obtained with 10 mM caffeine. When rates were biphasic, faster rates were considered. Variability was observed in the response to $20 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$: 10 fibres were fully responsive, 4 fibres were partially responsive ($\sim 20\%$ of the effect shown in **b**) and 3 fibres were not responsive at all. Such differences can be tentatively ascribed to variations in (1) $\text{Ins}(1,4,5)\text{P}_3$ hydrolysis (which should be already low at 0.1 mM free Mg^{2+})^{23,24}; (2) Ca^{2+} loading by the fibres and (3) fibre diameter and relative diffusion times into the fibre. In some experiments, we used $\text{Ins}(1,4,5)\text{P}_3$ derived from ox brain $\text{PtdIns}(4,5)\text{P}_2$ (ref. 30) and results were comparable.

Methods. Rabbit adductor muscle bundles were chemically skinned^{17,35} and stored at -20°C in a solution containing 50% glycerol, 5 mM $\text{K}_2\text{-EGTA}$, 0.17 M K-propionate pH 7.0, 2.5 mM $\text{K}_2\text{-Na}_2\text{-ATP}$, 2.5 mM Mg-propionate and 10 mM imidazole propionate. For isometric force measurements, segments of single fibres were positioned between two clamps, one of them attached to a strain gauge, and stretched to 130% of the slack length³⁵. After 30–45 s in solution R (5 mM $\text{K}_2\text{-EGTA}$, 0.17 M K-propionate , 2.5 mM Mg-propionate , 5 mM $\text{K}_2\text{-Na}_2\text{-ATP}$ and 10 mM imidazole propionate pH 7.0) fibres were allowed to accumulate Ca^{2+} for 30 s in a $p\text{Ca}$ 6.4 loading solution (0.17 M K-propionate , 2.5 mM Mg-propionate , 5 mM $\text{K}_2\text{-Na}_2\text{-ATP}$, 2.5 mM CaCl_2 , 5 mM $\text{K}_2\text{-EGTA}$ and 10 mM imidazole propionate pH 7.0), rinsed for 20–30 s in solution W (solution R without EGTA) and then challenged with $\text{Ins}(1,4,5)\text{P}_3$. Free Mg^{2+} (kept constant at 0.09 mM) and free Ca^{2+} ($0.40 \mu\text{M}$) were determined using association constants given in Orentlicher *et al.*³⁶. All experiments were carried out at room temperature ($22\text{--}24^\circ\text{C}$) and ionic strength 0.2 M.

efflux which, in turn, triggers a massive Ca^{2+} release by opening TC Ca^{2+} -gated Ca^{2+} channels¹³. The relevance of this model to *in vivo* Ca^{2+} release could not be assessed fully in the skinned fibre preparation because of an experimental constraint that cannot be overcome, that is, exogenous $\text{Ins}(1,4,5)\text{P}_3$ must be added to the bathing solution from where it diffuses into the fibre. Interestingly, $\text{Ins}(1,4,5)\text{P}_3$ triggered Ca^{2+} release from the SR of mechanically skinned frog skeletal muscle fibres²³, and $\text{Ins}(1,4,5)\text{P}_3$ was more effective²³ if its hydrolysis was inhibited by either low free Mg^{2+} (ref. 24) or 2, 3-diphospho-D-glycerate²⁴.

The proposed model is still highly speculative and several crucial questions remain unanswered: does $\text{PtdIns}(4,5)\text{P}_2$ phosphodiesterase exist in TT membranes and, more importantly, does polyphosphoinositide breakdown take place during muscle activation? In this respect we note that Novotny *et al.*²⁵ have shown that K^+ depolarization increased ^{32}P -labelling of phosphatidylinositol in frog sartorius muscles, and that the phosphatidylinositol response is caused by K^+ depolarization and not by the ensuing transient Ca^{2+} release from the SR. Moreover, preliminary data obtained by us indicate that water-soluble inositol phosphates could be extracted from rat diaphragm muscle and that $^3\text{H-myoinositol}$ incorporation into membrane phospholipids increased following a 5-s tetanus. Finally, in frog sartorius and semitendinosus muscles, a short-pulse tetanus followed by rapid freezing²⁶ increased the level of phosphatidylinositol, $\text{Ins}(4,5)\text{P}_2$ and $\text{Ins}(1,4,5)\text{P}_3$ to levels three to four times those in controls (J. Vergava and R. Y. Tsien, personal communication). Taken together, these observations suggest that

Table 1 Effect of ruthenium red on $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release from isolated SR fractions, and effect of $\text{Ins}(1,4)\text{P}_2$ and $\text{Ins}(4,5)\text{P}_2$

Additions	Extent of Ca^{2+} release (% of Ca^{2+} accumulated)	
	LSR	TC
$\text{Ins}(1,4,5)\text{P}_3$ ($25 \mu\text{M}$)	17	47
$\text{Ins}(1,4,5)\text{P}_3$ ($25 \mu\text{M}$) plus ruthenium red ($20 \mu\text{M}$)	7	8
$\text{Ins}(1,4)\text{P}_2$ ($25 \mu\text{M}$)	2	1
$\text{Ins}(4,5)\text{P}_2$ ($25 \mu\text{M}$)	3	1

Ca^{2+} uptake and Ca^{2+} release were measured as described in Fig. 1 legend. $\text{Ins}(1,4)\text{P}_2$ and $\text{Ins}(4,5)\text{P}_2$ were prepared by alkaline hydrolysis of ox brain $\text{PtdIns}(4,5)\text{P}_2$ (ref. 30). With no additions there was a basal Ca^{2+} release that corresponded to the slightly lower steady state attained after peak Ca^{2+} accumulation (compare Fig. 1a, dashed line), and accounted for approximately 4% and 8% of the Ca^{2+} accumulated in LSR and TC fractions, respectively. The extent of $\text{Ins}(1,4,5)\text{P}_3$ induced Ca^{2+} release was calculated as the difference between total and basal Ca^{2+} release (see Fig. 1). Data are from a typical experiment in which duplicate determinations were performed.

polyphosphoinositide metabolism may be causally related to muscle activation.

Another fundamental question is whether the phosphoinositide breakdown rate is fast enough (in the millisecond range) to exert a physiological role in skeletal muscle contraction. This question was not directly addressed. However, the recent findings^{27,28} that $\text{PtdIns}(4,5)\text{P}_2$ breakdown is involved in phototransduction, a quite complex and rapid event²⁹, encourage us not to rule out *a priori* such a mechanism for muscle activation.

We thank Dr R. F. Irvine for the gift of $\text{Ins}(1,4)\text{P}_2$, $\text{Ins}(4,5)\text{P}_2$ and a sample of specially purified $\text{Ins}(1,4,5)\text{P}_3$, and Dr S. Treves for critically reading the manuscript.

Received 21 January; accepted 3 May 1985.

- Porter, K. R. & Palade, G. E. *Biophys. biochem. Cytol.* **3**, 269-300 (1957).
- De Meis, L. in *Transport in Life Sciences* Vol. 2 (ed. Bitar, E. E.) 1-163 (Wiley, New York, 1981).
- Peachey, L. D. *J. Cell Biol.* **25**, 209-231 (1965).
- Franzini Armstrong, C. *J. Cell Biol.* **47**, 488-499 (1970).
- Huxley, A. F. & Taylor, R. E. *J. Physiol., Lond.* **144**, 426-441 (1958).
- Winegrad, S. *J. gen. Physiol.* **55**, 77-88 (1970).
- Endo, M. *Physiol. Rev.* **57**, 71-108 (1977).
- Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
- Suematsu, E., Hirata, M., Hashimoto, T. & Kuriyama, H. *Biochem. biophys. Res. Commun.* **120**, 481-485 (1984).
- Hirata, M., Suematsu, E., Hashimoto, T., Hamachi, T. & Koga, T. *Biochem. J.* **223**, 229-236 (1984).
- Saito, A., Seiler, S., Chu, A. & Fleischer, S. *J. Cell Biol.* **99**, 875-885 (1984).
- Campbell, K. P. & Shamoo, A. E. *J. Membrane Biol.* **54**, 73-80 (1980).
- Miyamoto, H. & Racker, E. *J. Membrane Biol.* **66**, 193-201 (1982).
- Mitchell, R. D., Volpe, P., Palade, P. & Fleischer, S. *J. biol. Chem.* **258**, 9867-9877 (1983).
- Kim, D. H., Ohnishi, S. T. & Ikemoto, N. *J. biol. Chem.* **258**, 2365-2374 (1983).
- Volpe, P., Palade, P., Costello, B., Mitchell, R. D. & Fleischer, S. *J. biol. Chem.* **258**, 12434-12442 (1983).
- Wood, D. S., Zollman, J. R., Reuben, J. P. & Brandt, P. W. *Science* **187**, 1075-1076 (1975).
- Julian, F. J. *J. Physiol., Lond.* **218**, 117-145 (1971).
- Somlyo, A. V., Gonzales-Serratos, H., Shuman, H., McLellan, G. & Somlyo, A. P. *J. Cell Biol.* **90**, 577-594 (1981).
- Weber, A. & Herz, R. *J. gen. Physiol.* **52**, 750-759 (1968).
- Lüttgau, H. C. & Oetlicher, H. *J. Physiol., Lond.* **194**, 51-74 (1968).
- Stephenson, E. W. *J. gen. Physiol.* **77**, 419-443 (1981).
- Vergara, J. & Tsien, R. Y. *Biophys. J.* **47**, 351a (1985).
- Downes, C. P., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169-177 (1982).
- Novotny, L., Saleh, F. & Novotna, R. *Gen. Physiol. Biophys.* **2**, 329-337 (1983).
- Ferenci, M. A., Homsher, E. & Trentham, D. R. *J. Physiol., Lond.* **352**, 575-599 (1984).
- Brown, J. E. *et al. Nature* **311**, 160-163 (1984).
- Ghalayini, A. & Anderson, R. E. *Biochem. biophys. Res. Commun.* **124**, 503-506 (1984).
- Gould, G. H. & Korenbrot, J. I. *Proc. natn. Acad. U.S.A.* **77**, 5557-5561 (1980).
- Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **221**, 269-272 (1984).
- Meissner, G. *Molec. cell. Biochem.* **55**, 62-82 (1983).
- Zorzato, F., Salviati, G., Facchinetti, T. & Volpe, P. *J. biol. Chem.* (in the press).
- Lowry, O. H., Rosebrough, N. J., Farr, A. R. & Randall, R. J. *J. biol. Chem.* **192**, 265-275 (1951).
- Downes, C. P. & Michell, R. H. *Biochem. J.* **133**-140 (1981).
- Salviati, G., Sorenson, M. M. & Eastwood, A. R. *J. gen. Physiol.* **79**, 603-632 (1982).
- Orentlicher, M., Brandt, P. W. & Reuben, J. P. *Am. J. Physiol.* **233**, C127-C134 (1977).

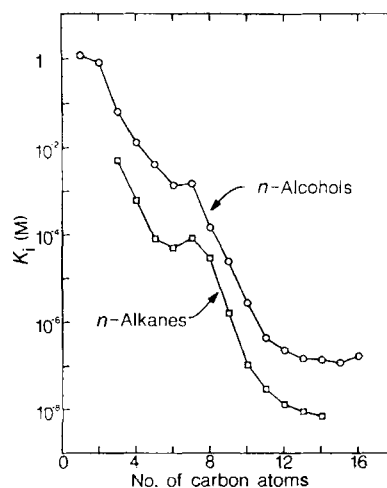


Fig. 1 Inhibition constants K_i for the homologous series of *n*-alcohols and *n*-alkanes acting on the enzyme luciferase. Firefly luciferase was purified, its activity assayed and the inhibition constants for competition with the natural substrate luciferin determined as described previously¹⁴. Standard errors in the K_i values were $\leq 20\%$. As before, the data were analysed using a simple binding model¹⁴ in which *N* anaesthetic molecules bind to the enzyme with equal affinities but only one molecule is sufficient to cause inhibition. We find that the break between *N* = 2 and *N* = 1 occurs for the alcohols between *C*₇ and *C*₈. For the alkanes, we confirmed that this break occurs between *C*₆ and *C*₁₀. However, the low activities of the intervening members meant that the exact point could not be ascertained; we took it as being the same as for the alcohols. For the largest molecules with *N* = 2, it seems likely that the two molecules bind with different affinities, although our data were not sufficiently precise to detect this. In such cases, the K_i values we have determined represent weighted averages. The alcohols *C*₁₁→*C*₁₆ were added as ethanolic solutions, such that the final ethanol concentration never exceeded 5% of the K_i for ethanol; the controls nonetheless contained the appropriate concentration of ethanol. Solutions containing the gaseous alkanes *C*₃ and *C*₄ were prepared by bubbling the gases through scintered glass; results obtained after 1 day were identical to those obtained after 2 days of bubbling. The final concentrations were calculated using the mole fraction solubilities given in ref. 18. Saturated solutions of alkanes *C*₅→*C*₁₃ were prepared by prolonged and vigorous shaking with small (~20 ml) volumes of buffer. In addition, alkanes *C*₈→*C*₁₄ were added as ethanolic solutions such that their final concentrations were only about five times in excess of their saturated solutions. Results obtained using these two methods gave excellent agreement. Values for the saturated aqueous solutions of these alkanes were calculated using the straight line relationship shown in Fig. 2b. For each anaesthetic, the concentration of enzyme was always kept well below the lowest concentration of anaesthetic used. The alcohols and alkanes were the purest grades available from Sigma or BDH.

Mapping of general anaesthetic target sites provides a molecular basis for cutoff effects

N. P. Franks & W. R. Lieb

Biophysics Section, Blackett Laboratory, Imperial College of Science and Technology, Prince Consort Road, London SW7 2BZ, UK

A longstanding and unresolved problem in general anaesthesia is the so-called 'cutoff' effect; as one ascends a homologous series of anaesthetic agents, the potencies progressively increase with anaesthetic size but then, rather suddenly, anaesthetic potency disappears¹⁻⁵. Curiously, this cutoff in potency occurs at very different points in different series. Various explanations have been offered^{6,7}, usually based on the notion that lipid bilayers are the primary target sites in general anaesthesia^{3,4,7}. However, accumulating evidence now suggests that proteins are the primary sites of action⁸⁻¹⁴. Here we demonstrate cutoff effects for the anaesthetic inhibition of a soluble protein (firefly luciferase) which mirror those found for general anaesthesia, and we describe how the molecular architecture of the binding site accounts for the different cutoffs in the different homologous series. We show that this

behaviour is a natural consequence of anaesthetics binding to an amphiphilic protein pocket of circumscribed dimensions. When general anaesthetic target sites in animals and the luciferase protein are mapped out using the fine details of the potency data, remarkable similarities are revealed. Our results thus suggest that the target sites in general anaesthesia are amphiphilic pockets on proteins.

The best-characterized examples of homologous series which display cutoffs in anaesthetic potency are the *n*-alcohols and *n*-alkanes. General anaesthetic activity disappears at different points in these two series; the alcohols^{1,4} cut off at about *C*₁₃ while the alkanes^{2,5} cut off between *C*₆ and *C*₁₀, the exact point being species-dependent. To elucidate the molecular basis for these observations, we have determined the inhibition constants for homologous series of both alcohols and alkanes acting on a pure soluble protein. This protein, firefly luciferase, is competitively inhibited by a wide range of general anaesthetics at concentrations which are essentially identical to those which produce general anaesthesia in animals¹⁴.

The inhibition constants K_i for the homologous series of

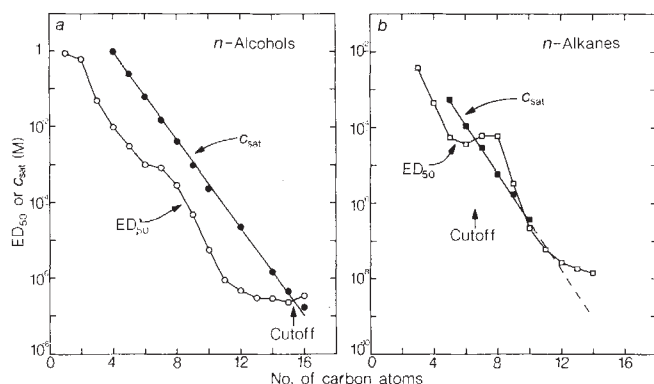


Fig. 2 The cutoff effect for the homologous series of *n*-alcohols (*a*) and *n*-alkanes (*b*). The ED_{50} values are the concentrations of anaesthetics which inhibited luciferase activity by 50% when the concentration of the substrate luciferin was equal to the K_m of the uninhibited enzyme ($14 \mu\text{M}$). Standard errors in the ED_{50} values were $\leq 20\%$. The c_{sat} values are the maximum possible aqueous concentrations (the solubilities) of the anaesthetics. When it was not possible to achieve 50% inhibition experimentally, ED_{50} values were estimated by extrapolation from the maximum attainable inhibitions. The c_{sat} values for the alcohols and alkanes are those compiled in ref. 19. The c_{sat} straight lines are drawn according to the equations given in ref. 19; for the alkanes we have assumed¹⁹ that the line can be extrapolated to C_{14} . Arrows indicate the positions of the abrupt cutoffs in potency.

n-alcohols and *n*-alkanes acting on firefly luciferase are plotted against the number of carbon atoms in Fig. 1. The two curves display some expected but also some surprising features. As might be expected for binding to a hydrophobic pocket, the alkanes bind more tightly than their corresponding alcohols, and the addition of a hydrophobic methylene group ($-\text{CH}_2-$) to either type of molecule tends to increase the binding affinity (that is, to reduce K_i). Interestingly, however, the alkanes only bind 5–50 times more tightly than their corresponding alcohols, very much less than the 10,000-fold factor for partitioning between water and a purely hydrophobic solvent such as hexadecane¹⁵. Evidently, the anaesthetic-binding site on luciferase behaves as if it is amphiphilic, containing polar as well as apolar parts. This might be expected of a binding site which presumably extends to the water/protein surface. What is more surprising is that there are two regions in each curve of Fig. 1, at about C_6 and after about C_{12} , where the K_i values level off. In other words, there are two regions in which making an anaesthetic molecule larger and more hydrophobic does not change the free energy of binding.

What molecular interpretation can be placed on these results? We have shown previously¹⁴ that the anaesthetic-binding site on luciferase has a volume roughly equal to that of two hexanol molecules and that for agents larger than heptanol only a single molecule can be accommodated. Thus, at around C_6 , when two molecules bind, the site is becoming fully occupied, so that additional CH_2 groups can no longer bind within the pocket but must remain in water, where they can make little or no contribution to the overall binding energy. The K_i curves therefore tend to level off in this region. Between C_8 and C_{12} , where only one molecule binds, additional CH_2 groups are no longer forced into a polar environment and thus contribute strongly to binding. After C_{12} , a single molecule is now large enough to fill the binding site so that additional CH_2 groups remain in water and the K_i curves level off once again. An alternative explanation for the levelling off at about C_6 is that the hydrocarbon chains encounter packing constraints or even a polar region near the bottom of the pocket, so that CH_2 groups in this region bind weakly. (Such a polar region has previously been postulated¹⁶ to account for the relatively weak binding of hydrocarbons to bovine serum albumin.) Although the details of the molecular interpretation must await future work, the levelling-off of the K_i curves can be accounted for by the anaesthetics binding to

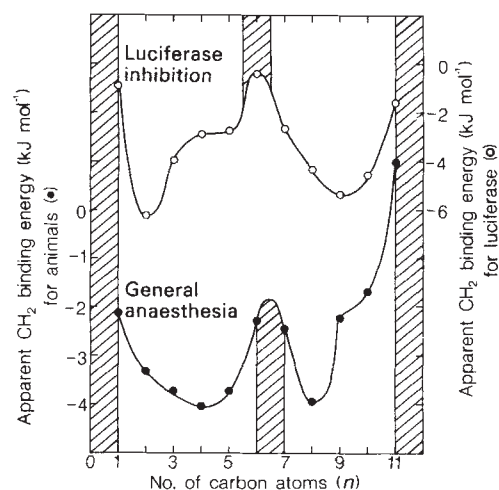


Fig. 3 Corresponding molecular features in the target sites in general anaesthesia and in the anaesthetic-binding pocket on the luciferase protein. The apparent changes in binding energy for a methylene group binding to luciferase (top) and target sites for general anaesthesia (bottom) are plotted as a function of chain length (n). Note the different scales in the vertical direction. Each point was calculated using the difference in thermodynamic free energy, $\Delta\Delta G^\ominus = -RT \ln[(ED_{50})_n / (ED_{50})_{n+1}]$ and alcohol ED_{50} values. The luciferase ED_{50} values are those plotted in Fig. 2*a* while those for general anaesthesia in animals (tadpoles) are taken from Tables 5 ($n = 1-8$) and 4 ($n = 9-12$) of ref. 20. True binding energies for animal target sites can only be calculated using inhibition constants which are, of course, unknown. So that a fair comparison between the data for general anaesthesia and luciferase inhibition could be made, we have used ED_{50} values throughout. The cross-hatched areas represent regions where methylene (CH_2) groups bind relatively weakly. A molecular interpretation of these features is given in the text.

an amphiphilic pocket of circumscribed dimensions.

To make a direct comparison with known ED_{50} concentrations for general anaesthesia, we determined luciferase ED_{50} concentrations for all of the above agents. (An ED_{50} concentration is one which produces a 50% effect, either anaesthetizing 50% of a population of animals or reducing luciferase activity by 50%.) Since alcohols larger than C_{11} all bind to the protein with roughly equal affinities (see Fig. 1), it might be expected that all of these large molecules would be equally effective at inhibiting enzyme activity. However, while C_{12} , for example, is among the most potent alcohols, we found that C_{16} was not an effective inhibitor. The resolution of this paradox is clear from the data shown in Fig. 2*a*, where we have plotted alcohol ED_{50} concentrations together with the corresponding maximum aqueous solubilities (c_{sat}). It can be seen that C_{16} is ineffective simply because its maximum achievable aqueous concentration lies well below that necessary to halve the enzyme activity.

Because of the similarities between the alkane and alcohol K_i curves in Fig. 1, one might suppose that the ED_{50} cutoff points for the two series would occur at similar places. However, as observed for general anaesthesia, we found that the alkanes cut off much earlier. The reason for this cutoff is clear from the data in Fig. 2*b*: the c_{sat} line intercepts the alkane ED_{50} curve between C_6 and C_7 . After C_8 , the site is occupied by only a single molecule and binding again increases such that C_{10} and C_{11} become active, but then, once more, the finite size of the pocket comes into play. (Although this recovery of activity has not, to our knowledge, been looked for in animals, it might well be peculiar to luciferase.)

Thus, for inhibition of luciferase activity, alcohol and alkane cutoffs appear at very different points. The alcohol cutoff at C_{16} is a consequence of the levelling-off of the ED_{50} curve (Fig. 2*a*) beyond about C_{12} , a result of the circumscribed dimensions of the binding pocket. The alkane cutoff at C_7 , on the other hand, results from the unexpected levelling-off of the ED_{50} curve (Fig. 2*b*) at about C_6 . In practical terms, the cutoffs occur not because

the anaesthetic molecules bind weakly to the protein but because the maximum aqueous solubilities of the inactive agents lie below the concentrations needed to half-inhibit activity. Our observations provide a plausible basis for interpreting cutoffs in anaesthetic potencies in animals. For tadpoles^{1,4}, the alcohol cutoff occurs slightly earlier (at C_{13} instead of C_{16}), and this may reflect a somewhat smaller volume for general anaesthetic target sites in animals. The alkane cutoffs for different animals^{2,5} occur between C_6 and C_{10} , similar to our value of C_7 .

In lipid theories, alcohol⁴ and alkane⁷ cutoffs have been accounted for in terms of an abrupt decrease in absorption into lipid bilayers. Very recently, however, it has been shown¹⁷ that partition coefficients for *n*-alkyltrimethylammonium ions continue to increase, at least up to C_{18} , with no suggestion of a sudden cutoff in absorption. It seems unlikely, therefore, that potency cutoffs in other amphiphiles, such as the *n*-alcohols, can be explained in terms of bilayer solubility.

Finally, in Fig. 3 we have used alcohol ED₅₀ data to make a direct comparison between anaesthetics binding to the luciferase protein and to the target sites involved in general anaesthesia, by plotting the apparent changes in CH₂ binding energies as the alcohol molecule increases in length. The cross-hatched areas indicate those regions on the luciferase and central nervous system target sites where CH₂ groups bind relatively weakly. The remarkable similarity between the data for general anaesthesia and for inhibition of the protein luciferase suggests that corresponding molecular features may be present in the target

sites for general anaesthesia in animals.

Although these target sites in animals have yet to be identified, a pocket of circumscribed dimensions and amphiphilic nature might be expected for a protein site which must selectively bind and orient relatively apolar ligands containing at least one polar group.

We thank Dr E. B. Smith and Professor D. M. Blow for stimulating discussions and the latter for helpful comments on the manuscript. This research was supported by grants from the MRC and the BOC Group, Inc.

Received 4 April; accepted 7 June 1985.

1. Meyer, K. H. & Hemmi, H. *Biochem. Z.* **277**, 39–71 (1935).
2. Gary-Bobo, C. & Lindenberg, B. A. *C. r. heb. Séanc. Acad. Sci., Paris* **234**, 2111–2113 (1952).
3. Mullins, L. J. *Chem. Rev.* **54**, 289–323 (1954).
4. Pringle, M. J., Brown, K. B. & Miller, K. W. *Molec. Pharmac.* **19**, 49–55 (1981).
5. Mullins, L. J. in *Alterations of Chemical Equilibrium in the Nervous System* (ed. Lajtha, A.) 395–421 (Plenum, New York, 1971).
6. Ferguson, J. *Proc. R. Soc. B* **127**, 387–404 (1939).
7. Haydon, D. A., Hendry, B. M., Levinson, S. R. & Requena, J. *Nature* **268**, 356–358 (1977).
8. Adey, G., Wardley-Smith, B. & White, D. *Life Sci.* **17**, 1849–1854 (1976).
9. Middleton, A. J. & Smith, E. B. *Proc. R. Soc. B* **193**, 173–190 (1976).
10. Franks, N. P. & Lieb, W. R. *Nature* **274**, 339–342 (1978).
11. Richards, C. D. *et al. Nature* **276**, 775–779 (1978).
12. Franks, N. P. & Lieb, W. R. *Nature* **300**, 487–493 (1982).
13. Smith, E. B. *et al. Nature* **311**, 56–57 (1984).
14. Franks, N. P. & Lieb, W. R. *Nature* **310**, 599–601 (1984).
15. Stein, W. D. *Transport and Diffusion Across Cell Membranes* (Academic, New York, 1986).
16. Tanford, C. *The Hydrophobic Effect* 2nd edn (Wiley, New York, 1980).
17. Requena, J. & Haydon, D. A. *Biochim. biophys. Acta* **814**, 191–194 (1985).
18. Wilhelm, E., Battino, R. & Wilcock, R. J. *Chem. Rev.* **77**, 219–262 (1977).
19. Bell, G. H. *Chem. Phys. Lipids* **10**, 1–10 (1973).
20. Brink, F. & Posternak, J. M. *J. cell. comp. Physiol.* **32**, 211–233 (1948).

Tissue-specific modulation of a set of related cell surface antigens in *Drosophila*

Michael Wilcox & Maria Leptin

Medical Research Council Laboratory of Molecular Biology,
Hills Road, Cambridge CB2 2QH, UK

To be integrated correctly into the developing organism, cells need to interact with an environment that includes neighbouring cells, the extracellular matrix and soluble factors. These interactions probably require specific recognition events mediated by cell-surface molecules, and evidence has accumulated for the presence of such molecules in several organisms (see, for example, refs 1–6). In *Drosophila*, a family of cell-surface glycoprotein complexes, the position-specific (PS) antigens, may be involved in such processes^{7–11}. PS antigens show spatially restricted distributions on imaginal disks that vary during development; their expression on any particular cell correlates with the position of that cell in the developing disk epithelium rather than with its ultimate differentiation. In contrast to the situation in disks, however, immunofluorescence revealed few regional restrictions of PS-antigen expression on other larval tissues^{7,11}. If the PS antigens have a general role in morphogenesis, then widespread regional variation in their distributions might be expected. We now report that, by using polyacrylamide gel electrophoresis (PAGE), a technique that gives higher resolution than immunofluorescence, we detect an increase in both the concentration and the diversity of the complexes during embryogenesis, beginning at gastrulation. Further, each larval and adult tissue examined carries its own characteristic set of complexes. This tissue- and region-specific diversity is achieved by combining different members of a set of related components with a glycoprotein common to all the complexes.

Monoclonal antibodies against the PS1 and PS2 antigens were defined by their preferential binding to cells of the dorsal and ventral compartments, respectively, of the mature wing imaginal disk^{7,8}. The antibodies precipitate different but related complexes of glycoproteins from disk lysates (refs 8, 9; lanes a in Fig. 1). Each antibody recognizes components unique to its own complex: glycoprotein 116 (gp116) and gp120 in the case of

PS1 antibodies; gp125 for the PS2 antibodies⁹. A third class of antibodies, PS3, recognize determinants on a gp110 component common to the complexes. PS3 antibodies thus precipitate both the PS1 and PS2 complexes together with at least one further complex containing a gp92 doublet.

To investigate the expression of the antigens on tissues other than imaginal disk, we analysed the composition of embryonic antigens at different stages of development (see Fig. 1 legend for a description of the timing of the main events of embryogenesis) and on different larval and adult tissues to determine whether we could detect variations or restricted distributions not revealed by immunofluorescence staining. Lysates of embryos and of first (24 and 28 h)- and second (52 h)-instar larvae were extracted sequentially, first with PS1, then with PS2 and finally with PS3 antibodies (see Fig. 1 legend) and the antigen components were separated by SDS-PAGE. The main findings, shown in Fig. 1, can be summarized as follows.

(1) The PS complexes are present only at low concentrations until after the onset of gastrulation (which begins ~3.5 h after egg laying, shortly after the formation of the cellular blastoderm). In Fig. 1, only the 1 h and 4 h time points are shown but sampling at other times within this period confirms the observation. From about 4 to 6 h, all PS antigens show a dramatic increase of both amount and complexity.

(2) PS1 and PS2 antibodies extract discrete subsets of the total PS complexes at each stage, leaving behind others which can be extracted subsequently by PS3 antibodies.

(3) Different components appear and reach their maximum levels at different times and several appear only transiently or show fluctuations in their level, for example, the PS1 component of relative molecular mass (M_r) 120,000 (120K), or the PS3 doublet at 92K.

(4) Several components found in embryonic PS antigens do not occur in the disk antigens (lanes a of Fig. 1). Thus, each antibody extracts different complexes from different lysates. Moreover, the large number of components often seen in the antigen extracted from embryos by any given antibody suggests that each antibody may be precipitating several different complexes, each expressing a molecule carrying the determinant recognized by that antibody, but containing different components attached to it. Extrapolating from what we know of the spatially restricted distribution of PS antigens on imaginal disks, it seemed likely that the additional complexes in the embryo

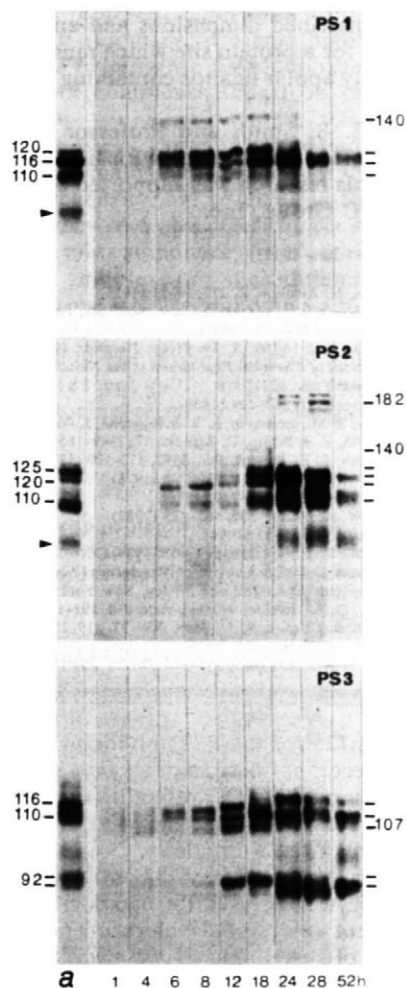
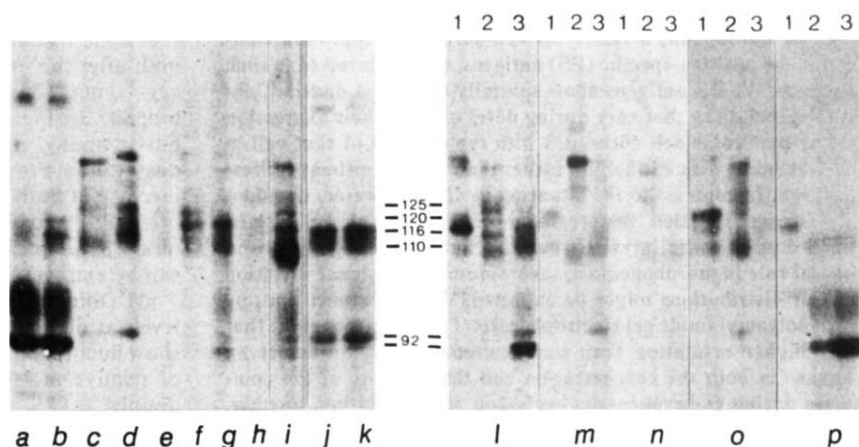


Fig. 1 Embryonic and early larval PS antigens. Embryos or larvae of the ages shown were lysed, the lysates extracted in turn with PS1, PS2 and finally PS3 antibodies, and the PS-antigen components resolved by SDS-PAGE. Preliminary experiments showed that the PS1 and PS2 antibody levels were sufficient to remove all antigens recognized by those antibodies. The overall pattern of components and the timing of their appearance is reproducible from experiment to experiment. Lanes *a* show, for comparison, the components of PS1, PS2 and PS3 antigens extracted sequentially from an imaginal disk lysate. The band indicated by an arrowhead at ~100K is a breakdown product of gp110 (ref. 9).

Methods. Eggs were collected from *Drosophila melanogaster* flies (Oregon R strain) over 1-h laying periods, dechorionated and fractured by sonication at the times shown (=time of development $\pm$ 30 min), and proteins were solubilized in 10 vol. of a non-ionic detergent buffer containing 20 mM Tris-HCl pH 8.1, 100 mM NaCl, 1 mM each of $MgCl_2$ and $CaCl_2$, 0.2% each of Nonidet-P40 and Triton X-100, 1% kallikrein inactivator and 1 mM phenylmethylsulphonyl fluoride. Equal volumes (50 μ l) of packed embryos or larvae were used at each time point. Each lysate was extracted in turn with PS1, then PS2 and finally PS3 antibodies attached to Affi-Gel 10 (10- μ l beads in each case). After washing, antigens were eluted from the antibody support columns with 10 μ l sample buffer (125 mM Tris-HCl pH 6.5, 2% SDS, 10% glycerol), reduced by addition of 2-mercaptoethanol to 5% v/v and components resolved by SDS-PAGE on thin (0.75 mm) 8% gels. The components are all glycoproteins and were visualized by the procedure of Hawkes¹³: nitrocellulose (Western) blots¹⁴ from gels are incubated successively in concanavalin A (Con A) and horseradish peroxidase (HRP), and HRP activity is then detected in a staining reaction. M_r s were determined using 29–205K protein markers (Sigma). All procedures have been described previously^{8,9}, as has the isolation of the antibodies. Each PS antibody class contains several independently isolated monoclonal antibodies with identical specificities^{8,9}. *D. melanogaster* embryogenesis lasts for ~22 h at 25°C. The timing of some of the major events is as follows (for review see ref. 12): the cellular blastoderm forms at 2.5–3 h after egg laying, gastrulation begins at 3.5 h, morphologically visible segmentation at ~7 h. Most tissue rudiments become visible between 3 and 10 h and organogenesis continues until shortly before the first-instar larva hatches at 22 h. A moult 24 h later marks the beginning of the second larval instar, which lasts until the second moult at ~72 h after egg laying. The third larval stage ends at 120 h, when the animal pupates and begins metamorphosis.

Fig. 2 PS antigens from larval and adult tissues. Lanes *a*–*i*, total PS antigens extracted by PS3 antibodies. Subsequent extraction of the lysates with PS1 and PS2 antibodies yielded no further antigens, showing that all complexes contained the common gp110 component recognized by PS3 antibodies. Lanes 1–3 of *l*–*p* are antigens extracted from lysates by PS1, PS2 and PS3 antibodies in sequence (see Fig. 1 legend). Tissues: *a* and *p*, larval gut; *b*, adult gut; *c*, adult testis; *d* and *m*, larval salivary gland; *e* and *n*, larval brain; *f* and *o*, adult ovary (anterior half); *g*, oocytes; *h*, larval fat body; *i*, larval cuticle (contains epidermis plus attached musculature); *j* and *k*, imaginal disk with and without larval gut lysate, respectively (see below); *l*, larval head (contains most imaginal disks together with larval foregut, brain and some epidermis). In all these anti-



gens, the band in the region of 110K (and often one at ~100K) is recognized by the monoclonal antibody DX4C8 (an antibody that binds to disk antigen gp110 and its 100K breakdown product on Western blots, but not on cells or in solution⁹; data not shown).

Methods. Tissues were dissected out of mature third-instar larvae or young adults of the *D. melanogaster* Barton strain (very similar results were obtained using tissues from Oregon R or Canton S animals), lysed by sonication in non-ionic detergent buffer (see Fig. 1 legend) and antigens extracted by the Affi-Gel 10-bound PS antibodies described above. Equivalent volumes (~50 μ l) of packed tissue were used in each case (this volume required variable numbers of organs: 75 pairs of larval salivary glands and 40 adult testes, for example). Antigen components were resolved by SDS-PAGE, then (with the exception of lanes *j* and *k*) transferred to nitrocellulose and visualized with Con A/HRP (see Fig. 1 legend). *j*, *k*, PS3 antigens were extracted from a lysate of ³⁵S-methionine-labelled imaginal disks¹⁵ after incubation for 2 h at 24°C either alone (*j*) or with the lysate from 50 μ l larval gut tissue (*k*) (in another experiment a similar antigen pattern was obtained after incubation with intact extirpated larval gut). ³⁵S-labelled products were visualized by fluorography¹⁶.

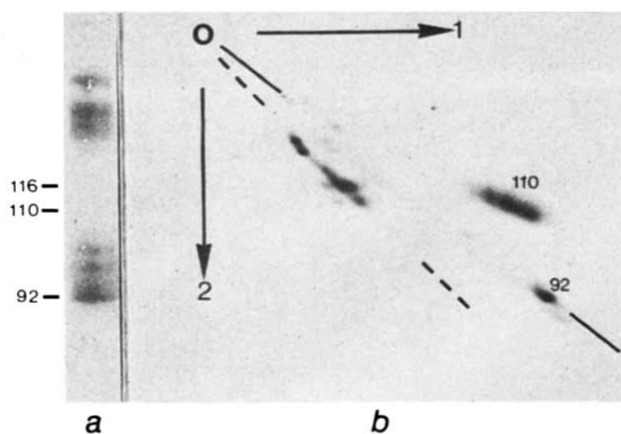


Fig. 3 Two-dimensional non-reducing/reducing SDS-PAGE of late embryo PS antigens. Components of total PS antigen extracted from a lysate of 18-h embryos were first fractionated by non-reducing SDS-PAGE on a thin (0.75 mm) 8% gel (*a*). Note that all PS components have moved away from the 110–120K region, some to higher, others to lower apparent M_r s. A lane identical to that shown in *a* was cut out of the gel, treated for 1 h at room temperature with sample buffer containing 5% 2-mercaptoethanol, then placed on top of a second, slightly thicker (0.9 mm) 8% gel and the reduced proteins re-electrophoresed (*b*). (For details of this procedure see ref. 9.) Products were transferred to nitrocellulose and visualized by Con A/HRP staining. Gp110, which migrates more rapidly in the non-moving dimension, lies above the diagonal (solid line) of the gel; three components can be resolved in this region (all of which are recognized by antibody DX4C8). Several components of higher M_r were retarded similarly in the first dimension (by ~15–20K) and so, in the two-dimensional gel, all lie below the diagonal on a line that also extrapolates to the origin (broken lines). This behaviour is reminiscent of that of gp116, 120 and 125 of the disk antigens, which contain similar polypeptides⁹. The gp92 doublet glycoproteins, which run identically in each dimension, lie on the diagonal.

might also be regionalized, perhaps localized to entire organs and tissues.

To investigate this possibility, we analysed antigens extracted by PS antibodies from lysates of several larval and adult tissues. Figure 2 shows the SDS gel patterns of total PS antigens (precipitated by PS3 antibodies). As expected, the complexes precipitated by the same antibodies from different tissues differ in their composition, with only the common gp110 occurring in all the antigens. These differences were highlighted in experiments in which tissue lysates were sequentially extracted with PS1, PS2 and PS3 antibodies. For example, larval brain (Fig. 2*n*, lanes 1–3) exhibits only low levels of an antigen that is almost exclusively PS2; in particular, little or no PS1 antigen is detectable, in agreement with immunofluorescence staining results¹¹. In the adult ovary (Fig. 2*o*, lanes 1–3), the population of antigens is composed almost entirely of the PS1 and PS2 subsets, while in the larval gut (Fig. 2*p*, lanes 1–3) complexes recognized only by PS3 antibodies predominate.

It is critical to exclude the possibility that this heterogeneity is an artefact; the reproducibility of the patterns is strong evidence against this. Furthermore, qualitatively identical patterns were obtained with tissues from a number of different fly strains (data not shown). Protease inhibitors were included in all preparations to minimize any effect of a differential occurrence of proteases in different tissues and several experiments were carried out to rule out this possible cause of artefacts. Gut PS antigens contain mainly components of low M_r , which might be candidates for degradation products, and the gut is most likely to contain proteolytic enzymes. Therefore, we incubated ³⁵S-methionine-labelled imaginal disk lysates with gut extracts or intact extirpated gut, before extraction with PS3 antibody. The ³⁵S-labelled PS3 antigen from this preparation was identical to that extracted from control lysates (Fig. 2*j*, *k*). Thus, we

could find no evidence of proteolytic or indeed any other modifying activity and conclude that the differences in the SDS gel patterns between different tissues represent differences in the composition of the complexes. In summary, it seems highly likely that the complex embryonic patterns we have described represent the sum of the different complexes found on the various organs and tissues present.

How are the PS antigens on the various organs, tissues and regions of the organism related? The main characteristics of the antigens, their ubiquity and heterogeneity, are reflected in the form of the complexes: each consists of a characteristic set of variable components together with one component, gp110, common to all of them. The fact that complexes of such different composition, for example, those from salivary gland and adult ovary (lanes 2 of *m*, *o* in Fig. 2), are precipitated by the same antibody, shows that at least some of the variable components are related. All these components (except gp92) migrate abnormally on two-dimensional non-reducing/reducing gels; they all appear on a single line extrapolating back to the origin but running below the gel diagonal. For example, in the case of the total PS antigens from 18-h embryos, six components ranging from 112K to ~140K lie on this line below the diagonal (Fig. 3*b*). The high- M_r components (>110K) of antigens isolated at other stages of development behave similarly (data not shown). Interestingly, the variable components of the disk antigens (gp116–125) show the same anomalous behaviour⁹, and partial proteolysis studies have indicated that these components contain similar polypeptides⁹ (while gp92 and gp110 are not related either to each other or to the other components). It seems likely, therefore, that at least some of the variable components of the embryonic and larval antigen complexes are related both to the variable components of the disk complexes and to one another. These variable components may be products of related genes or of one or a small number of common precursors (transcripts or polypeptides), which are modified according to position and time to give rise to the different tissue-specific components. Preliminary deglycosylation studies suggest that the variations do not stem from differential *N*-glycosylation (M.L. & M.W., unpublished observations).

The relatively indiscriminate expression of PS antigens on the majority of *Drosophila* tissues, as detected by immunofluorescence staining, turns out therefore to be only a superficial view of a more complicated picture. At the finer level of analysis used here, temporal and spatial differences in the composition and distribution of the PS antigens become apparent, and we begin to resolve the heterogeneity of the antigens and their specificity for organs and tissues. In effect, the network of position-specific antigens provides different tissues and regions of the organism with characteristic signatures distinguishing them from their surrounding tissues. It remains to be shown whether and how these signatures are used in the organization of the developing fly; the notion that they are in fact used rests largely on the correlation of the temporal and spatial changes of antigen expression with defined morphogenetic events. The restricted distributions of the antigens on imaginal disks and the changes in expression during development have been discussed previously^{7,8,10,11}. We now know also that the first appearance of the antigens during embryogenesis correlates closely with the onset of morphogenetic processes at gastrulation and that PS-antigen complexity increases dramatically during the ensuing hours when the regional organization of the embryo and the subsequent organogenesis are occurring. Our contention that this complexity results from the summing of the components of different tissue- and region-specific complexes first expressed in these early stages, is supported by the tissue-, organ- and region-specific antigen distributions found later in development.

While a role for the PS antigens in specific cell recognition is attractive, the known properties of the molecules may well be consistent with several other possible functions. The use of monoclonal antibodies specific for isolated variable components should help to elucidate the relationships between the antigens, and the location of their genes and subsequent exploitation of

Drosophila genetics should lead to the isolation of mutants that will allow us to establish the function of the antigens.

We thank Richard Smith for technical help, Danny L. Brower for the PS2 antibody (CF2C7) and Nick Brown and many of our colleagues for criticism of the manuscript. M.L. is the recipient of an EMBO long-term fellowship.

Received 27 February; accepted 9 May 1985.

1. Kemler, R., Babinet, C., Eisen, H. & Jacob, F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4449-4452 (1977).
2. Gerisch, G. in *Current Topics in Developmental Biology* Vol. 14, Pt 2 (eds Moscona, A. A. & Monroy, A.) 243-269 (Academic, New York, 1980).
3. Edelman, G. M. *Science* **219**, 450-457 (1983).
4. Damsky, C. H., Richa, J., Solter, D., Knudsen, K. & Buck, C. A. *Cell* **34**, 455-466 (1983).
5. Kruse, J. *et al. Nature* **311**, 153-155 (1984).
6. Rutishauser, U. *Nature* **310**, 549-553 (1984).
7. Wilcox, M., Brower, D. L. & Smith, R. J. *Cell* **25**, 159-164 (1981).
8. Brower, D. L., Wilcox, M., Piovant, M., Smith, R. J. & Reger, L. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7485-7489 (1984).
9. Wilcox, M., Brown, N., Piovant, M., Smith, R. J. & White, R. A. H. *EMBO J.* **3**, 2307-2313 (1984).
10. Brower, D. L. *Nature* **310**, 496-498 (1984).
11. Brower, D. L., Piovant, M. & Reger, L. A. *Dev. Biol.* **108**, 120-130 (1985).
12. Fullilove, S. L. & Jacobson, A. G. in *The Genetics and Biology of Drosophila* Vol. 2c (eds Ashburner, M. & Wright, T. R. F.) 105-227 (Academic, New York, 1978).
13. Hawkes, R. *Analyt. Biochem.* **123**, 143-146 (1982).
14. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
15. Cheney, C. M. & Shearn, A. *Dev. Biol.* **95**, 325-330 (1983).
16. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83-88 (1974).

Specific targeting of cytotoxic T cells by anti-T3 linked to anti-target cell antibody

Pilar Perez, Robert W. Hoffman, Stephen Shaw, Jeffrey A. Bluestone & David M. Segal*

Immunology Branch, National Cancer Institute, Bethesda, Maryland 20205, USA

The specificity of cytotoxic T lymphocytes (T_c) cells is conferred by an antigen-specific receptor, T_i , which in humans is physically associated with an invariant cell-surface glycoprotein, T3 (refs. 1-3). Monoclonal antibodies specific for either T3 and T_i are able to elicit a variety of T-cell responses such as lymphokine production, mitogenesis and cytotoxicity¹⁻¹¹. For example, human T_c cells lyse anti-T3-expressing hybridoma cells, but not cells of other specificity, presumably because anti-T3 on the hybridoma cells binds to T3 on the T_c cells and triggers lysis¹². Here, we have adapted approaches used in a different cytotoxic effector system, antibody-dependent cellular cytotoxicity (ADCC), to alter the specificity of T_c cell. Studies of ADCC¹³ showed that heteroaggregates containing anti-Fc receptor ($Fc\gamma R$) antibody cross-linked to a second antibody bind to $Fc\gamma R$ on ADCC effectors and cause them to kill target cells bearing antigen recognized by the second antibody. The present studies use anti-T3-containing heteroaggregates to re-target human T_c cells to cells for which we have appropriate antibodies, including xenogeneic tumour cells and chicken erythrocytes. These results extend previous observations on the role of T3 in triggering cytotoxicity and suggest that effector cell re-targeting could be used for *in vivo* treatment of neoplasms and other pathogens that express distinctive surface antigens.

Four different human anti-HLA-DPw2 (SB2) cytotoxic T-cell clones, 8.2, 8.4, 8.5 and 8.9 (ref. 14), were used in this study. All four express the T3 and T4 antigens and target-cell lysis mediated by three of the clones (8.4 being the exception) is inhibited by anti-T3 antibody¹⁴. Clones 8.9 and 8.5 were negative for $Fc\gamma R$ (as determined by flow cytometry¹⁵) and therefore cannot mediate ADCC (clones 8.2 and 8.4 were not tested).

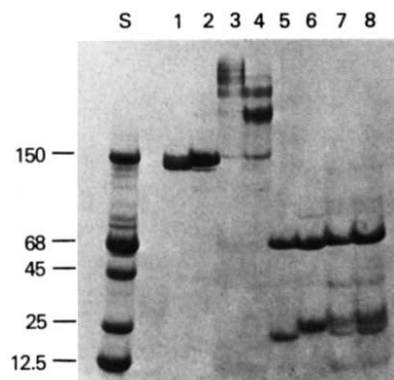


Fig. 1 SDS-polyacrylamide gel electrophoresis of cross-linked antibodies. Lanes 1-4, non-reduced; lanes 5-8, reduced. Lanes 1, 5, OKT3 (anti-T3); lanes 2, 6, 36-7-5 (anti- K^k); lanes 3, 7, OKT3 cross-linked to 36-7-5 (anti-T3 $\times$ anti- K^k), heavier fraction containing mainly trimers and larger oligomers; lanes 4, 8, lighter fraction of cross-linked material, containing mainly dimers and trimers. The cross-linked material reduces to light and heavy chains because a reducible cross-linking reagent was used in their preparation. Note that the cross-linked preparations contain light chains from both antibodies, as expected. Lane S, standards: IgG (relative molecular mass (M_r) 150,000), bovine serum albumin (68,000), ovalbumin (45,000), chymotrypsinogen (25,000) and cytochrome c (12,500).

Methods. OKT3 (Ortho) and 36-7-5²⁰ antibodies were purified from ascites by ion-exchange chromatography and gel filtration²¹. Each protein (5 mg) was incubated with a threefold molar excess of N-succinimidyl-3-(2-pyridyldithiol) propionate (Pharmacia). OKT3 was then reduced and coupled to 36-7-5 as described previously¹³. The hetero-cross-linked antibodies were eluted on a 1.6×90 -cm Ultrogel AcA 22 (LKB) column, fractions taken at relative elution volume $K_d = 0.08-0.33$ (heavy fraction) and $0.38-0.54$ (lighter fraction). Monomeric IgG elutes at (K_d) $0.63-0.71$. Samples were concentrated with a CX10 immiscible membrane (Millipore), and aliquots were denatured and run on a 2-16% gradient slab gel (Pharmacia). Other antibodies used were affinity-purified rabbit anti-DNP antibodies²² and anti-T4 monoclonal antibody^{1,3} (Ortho) purified as described above. In cross-linking anti-T3 to anti-DNP, an eightfold molar excess of SPDP to each antibody was used. This gave larger heteroaggregates than the threefold excess used in the other combinations, but also partially inactivated the antibodies, as shown by the inability of anti-T3 $\times$ anti-DNP to inhibit lysis of M16 cells by clone 8.9 (Table 2).

Cross-linked antibodies were prepared as described in Fig. 1 legend. Two cross-linked fractions were separated from monomeric IgG by gel filtration (Fig. 1). Preliminary experiments showed that both were equally active and they were therefore used interchangeably. In the present study, hetero-cross-linked antibodies are designated 'antibody 1 $\times$ antibody 2', for example, anti-T3 $\times$ anti- K^k .

The four clones were tested for their ability to lyse M16, a human lymphoblastoid cell line bearing the HLA-DPw2 antigen, RDM4, an H-2^k mouse tumour line, and EL-4, an H-2^b mouse tumour line (Table 1). As expected, these clones lyse the M16 target but not the murine tumour cells. Moreover, lysis of M16 by clones 8.9, 8.2 and 8.5 was blocked by anti-T3 $\times$ anti- K^k , showing that anti-T3 remained active in the heteroaggregate. The main finding of Table 1 is that anti-T3 $\times$ anti- K^k causes each of the T_c -cell clones to lyse the K^k -positive RDM4 targets, but not the K^b -expressing EL-4 cells. RDM4 was also lysed if either effector or target cells were preincubated with the heteroaggregates and washed before the experiment (data not shown).

These results suggest that anti-T3 $\times$ anti- K^k cross-links K^k on RDM4 to the T3 portion of the T-cell receptor complex on the T_c -cell clones and triggers the T_c cells to lyse the RDM4 targets. Figure 2 confirms specific intercellular cross-linking, as reflected in the formation of multicellular conjugates between effectors and targets. Figure 2a, b shows that in the absence of the

* To whom correspondence should be addressed at: Bldg 10, Rm 4B17, NIH, 9000 Rockville Pike, Bethesda, Maryland 20205, USA.

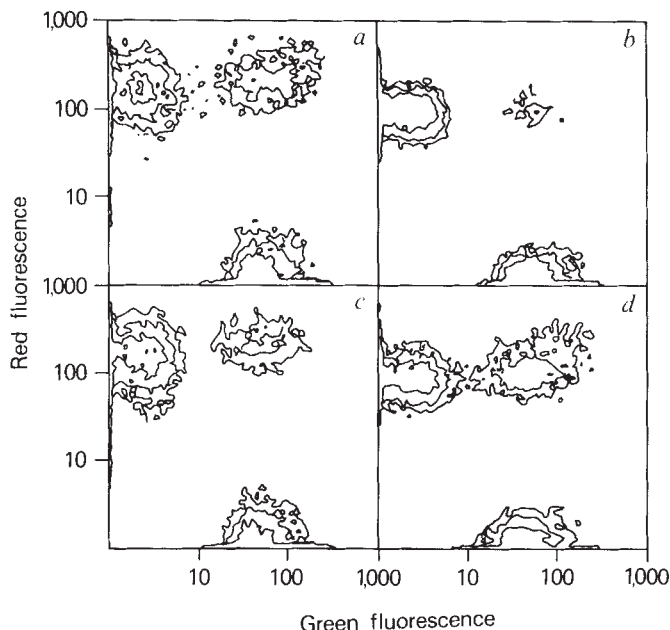


Fig. 2 Conjugate formation between T_c -cell clone 8.9 and tumour lines M16 (a, c) and RDM4 (b, d) in the absence (a, b) or presence (c, d) of anti-T3 $\times$ anti- K^k . Clone 8.9 cells were labelled with fluorescein isothiocyanate as described elsewhere²³. M16 and RDM4 cells (10^7 per ml) were incubated for 10 min at 37 °C with $10 \mu\text{g ml}^{-1}$ *N*-hydroxysuccinimidobiotin (Sigma) in PBS, pH 8.0, washed and incubated with phycoerythrin-avidin (20 μl of an appropriate dilution of *R*-phycoerythrin-avidin was added to 10^7 cells in 100 μl of PBS, pH 7.2, and incubated for 30 min at 24 °C). Conjugates were formed by mixing 8.9 cells with M16 or RDM4 cells at a 1:1 ratio, centrifuging and incubating for 20 min at 24 °C. In c and d, effectors were preincubated for 30 min at 0 °C with $10 \mu\text{g ml}^{-1}$ anti-T3 $\times$ anti- K^k before adding the target cells. After the 20-min incubation, cells were gently resuspended with a Pasteur pipette in 1 ml PBS at 0 °C and analysed with a Becton-Dickinson FACS analyser. Free 8.9 cells appear green only, free target cells appear red only and conjugates appear both red and green. 2×10^4 particles were analysed in each sample and contours were drawn through points containing 10, 20 and 40 particles. The percentages of effectors in conjugates were: a, 28%; b, 8%; c, 25%; d, 31%.

heteroaggregate, clone 8.9 forms the expected large number of conjugates with M16, but few conjugates with RDM4. By contrast, in the presence of heteroaggregates, the number of RDM4-8.9 conjugates increases dramatically while the number of M16-8.9 conjugates remains relatively unaffected. The latter observation is consistent with previous studies^{16,17} showing that anti-T3 blocks T_c -cell mediated lysis at a step following conjugate formation.

To test the specificities required for cross-linked antibodies to mediate lysis, we coupled anti-T3 to anti-dinitrophenyl(DNP), anti-T4 to anti- K^k and anti-DNP to anti- K^k , using procedures similar to those described for Fig. 1. Table 2 shows that the anti-T3 $\times$ anti-DNP confers anti-DNP specificity to T_c cell clones; in the presence of this heteroaggregate, T_c cells lyse TNP-RDM4 and TNP-EL4, but not unmodified EL4 or RDM4. Moreover, DNP hapten in solution strongly inhibits lysis of the haptenated targets. Interestingly, T_c cells will even lyse hapten-modified chicken red blood cells (CRBC) in the presence of anti-T3 $\times$ anti-DNP. Thus, the antibody heteroaggregates seem to override the normal requirement for lysis that T_c -cell receptors bind to molecules of the major histocompatibility complex (MHC) on target cells, as it is unlikely that either CRBC or the murine T_c -cell tumour lines EL4 and RDM4 express MHC molecules, which could be recognized by the class II-specific T_c -cell clones used here.

We next investigated whether anti-T3 must be a component of the antibody heteroaggregate, or whether monoclonal antibodies against any effector cell-surface molecule would trigger lysis. For this, we tested the effect of anti-T4 $\times$ anti- K^k on lysis

Table 1 Anti-T3 $\times$ anti- K^k heteroaggregates re-target T_c cells to K^k -expressing cells

Clone	Anti-T3 $\times$ anti- K^k	% Specific lysis of target cell		
		M16	RDM4	EL-4
8.9	—	47.0	6.6	0.8
8.9	+	11.5	53.3	0.2
8.2	—	44.9	9.1	3.3
8.2	+	20.4	43.9	—0.7
8.4	—	49.8	7.2	4.3
8.4	+	54.1	32.9	0.8
8.5	—	35.7	1.5	0.6
8.5	+	8.0	30.0	0.4

Lysis was measured in the presence or absence of $5 \mu\text{g ml}^{-1}$ anti-T3 $\times$ anti- K^k at a 10:1 effector: target ratio in a 3-h ^{51}Cr -release assay. Antibody was preincubated with effectors and targets for 30 min at 0 °C, before warming to 37 °C. Human T_c -cell clones 8.2, 8.4, 8.5 and 8.9 are specific for the HLA-DPw2 antigen^{14,18}. M16 (ref. 19) is a human Epstein-Barr virus-transformed lymphoblastoid cell line that expresses HLA-DPw1 and -DPw2 antigens. RDM4 and EL4 are H-2^k and H-2^b murine T_c -cell tumour lines. Per cent lysis was calculated as described elsewhere¹³, and spontaneous release was 8.9, 5.2 and 6.8% for M16, RDM4 and EL-4, respectively.

of M16 and RDM4. Results from experiments similar to those shown in Fig. 2 confirmed that anti-T4 $\times$ anti- K^k caused conjugates to form between clone 8.9 and RDM4 (21% of effectors were conjugated in the presence of heteroaggregate, 9% in its absence). However, this cross-linked preparation neither promoted lysis of RDM4 nor significantly blocked lysis of M16 by clone 8.9 (it has previously been demonstrated¹⁸ that anti-T4 does not inhibit lysis mediated by clone 8.9) (Table 2). In a second experiment (not shown in Table 2), T_c -cell clone 8.9 was treated with 1.5 mM trinitrobenzene sulphonate and tested for lysis against RDM4 in the presence of anti-DNP $\times$ anti- K^k . Control experiments showed that anti-DNP $\times$ anti- K^k promoted specific conjugate formation between TNP-8.9 and RDM4 and that the TNP-8.9 cells were able to lyse M16. However, TNP-8.9 cells did not lyse RDM4 cells in the presence of anti-DNP $\times$ anti- K^k (<2% lysis over a concentration range of 5–40 $\mu\text{g ml}^{-1}$ heteroaggregate). Therefore, cross-linking of target cells specifically to the T4 molecule and to a presumably large number of TNP-modified components on the effector cells did not lead to lysis. These data suggest that the cross-linking of target cells directly to the T_c -cell receptor complex is a requirement for lysis. In support of this view, earlier experiments¹² showed that several hybridoma cells expressing antibodies against effector cell-surface components other than T3 are not lysed by T_c -cell clones. In addition, Kranz *et al.*¹¹ have recently shown that a mouse T_c -cell clone will lyse targets to which anti-Ti antibodies were attached, but not targets bearing antibodies that bind to other effector cell-surface components.

From current results, there seems to be a striking parallel in the way in which ADCC effector cells and T_c cells lyse target cells; in both cases, direct attachment of the target cells to the appropriate receptor (Fc γ R on ADCC¹³ effectors or T_c -cell receptor complex on T_c cells) leads to target cell lysis, and attachment to other cell-surface components on the effector cell does not. These data strongly suggest that the target-specific receptors in both systems not only mediate conjugate formation, but also have an essential role in the post-conjugate phase of the lytic process. T3 is one cell-surface component that is clearly involved in the post-conjugate phase of T_c -cell mediated lysis, because anti-T3 blocks lysis but does not inhibit conjugate formation. Nevertheless, binding of the target cell directly to T3 results in target cell death. This suggests that the local aggregation of T3 at the cell/cell interface may be instrumental in triggering lysis, and that soluble anti-T3 blocks normal T_c -cell lysis by preventing local T3 aggregation.

We have demonstrated here that the specificities of T_c cells can be altered *in vitro* simply by treating them with antibody heteroaggregates. In fact, we have now been able to direct T_c

Table 2 Specificity of T_c-cell lysis in the presence of hetero-cross-linked antibodies

Target	% Specific lysis of target cells by clone 8.9					
	Antibody added					
	None	Anti-T3 × anti-K ^k	Anti-T3 × anti-K ^k + anti-T3	Anti-T3 × anti-DNP	Anti-T3 × anti-DNP + DNP	Anti-T4 × anti-K ^k
M16	37.8	-7.0	-29.7	22.3	—	29.9
RDM4	-8.7	38.2	2.0	-8.6	—	2.1
EL-4	-4.9	-4.2	—	-4.3	—	—
CRBC	1.1	—	—	1.3	0	—
M16-TNP	55.1	—	—	32.3	43.8	—
RDM4-TNP	-8.1	—	—	40.5	7.4	—
EL-4-TNP	0.4	—	—	49.1	0.3	—
CRBC-TNP	0.8	—	—	63.9	0	—

Experiments were performed as described in Table 1 legend using clone 8.9 cells as effectors. Clone 8.5 also lysed TNP-CRBC in the presence of anti-T3 × anti-DNP in a hapten-specific manner (data not shown). In all experiments shown, 10 µg ml⁻¹ cross-linked antibody was used, but 20 and 40 µg ml⁻¹ of anti-T4 × anti-K^k were also tested, with results similar to those shown here. Anti-T3 (50 µg ml⁻¹, column 4) and DNP-ε-aminocaproate (0.1 mM, column 6) were used as inhibitors. DNP-ε-aminocaproate did not inhibit lysis of M16 by clone 8.9 (data not shown). Target cells were modified with TNP by incubating them for 10 min at 37 °C in phosphate-buffered saline (PBS), pH 7.4, containing 3 mM trinitrobenzene sulphate. (—) Indicates experiment not done. Negative values for per cent lysis indicate that chromium release was less than spontaneous release. In the test sample, spontaneous release values for the various targets were: M16, 28.6%; RDM4, 11.4%; EL-4, 12.4%; CRBC, 2.9%; M16-TNP, 7.4%; RDM4-TNP, 15.4%; EL-4-TNP, 4.5%; and CRBC-TNP, 3.0%.

cells to lyse every target cell for which we have had an appropriate anti-target cell antibody. Based on these observations, it is reasonable to propose that T_c cells could be directed against various pathogenic cells *in vivo*, for example against tumour cells. By using anti-T3-containing heteroaggregates, presumably all T_c cells, and not a limited number of clonotypes, would be targeted to the tumour. In addition, such heteroaggregates could also target T-helper cells to a tumour (or other lesion) where they could serve as a source of lymphokines. Finally, appropriate heteroaggregates could be used to target both T cells (using anti-T3-containing heteroaggregates) and macrophages, granulocytes and natural killer cells (using anti-FcγR heteroaggregates) to the same lesion. Such strategies might greatly augment the normal immune response against tumours and other pathogenic cells, and could also provide insight into the normal role of various effector cells in combating infections and neoplasms.

We thank Drs Alfred Singer, Pierre Henkart and Ralph Quinones for helpful suggestions during the preparation of the manuscript, Ms Julie Titus for help in isolating antibodies and Ms Judy Kress for typing the manuscript. P. P. is the recipient of a Fulbright fellowship awarded by the Spanish Department of Education and Culture.

Note added in proof: Staerz *et al.*²⁴ have recently shown that anti-Ti-containing heteroaggregates will re-target murine T_c-cells.

Received 17 April; accepted 21 May 1985.

- Meuer, S. C., Acuto, O., Hercend, T., Schlossman, S. F. & Reinherz, E. L. *A. Rev. Immun.* **2**, 23-50 (1984).
- Haskins, K., Kappler, J. & Marrack, P. *A. Rev. Immun.* **2**, 51-66 (1984).
- Reinherz, E. L., Meuer, S. C. & Schlossman, S. F. *Immun. Rev.* **74**, 83-112 (1983).
- Lancki, D. W., Ma, D. I., Havran, W. L. & Fitch, F. W. *Immun. Rev.* **81**, 65-94 (1984).
- von Wussow, P., Platsoukas, C. D., Wiranowska-Stewart, M. & Stewart, W. E. *J. Immun.* **127**, 1197-1200 (1981).
- Meuer, S. C. *et al. J. exp. Med.* **158**, 988-993 (1983).
- van Wauwe, J. P., De May, J. R. & Goossens, J. G. *J. Immun.* **124**, 2708-2713 (1980).
- Tax, W. J. M., Hermes, F. F. M., Willems, H. W., Capel, P. J. A. & Koene, R. A. P. *J. Immun.* **133**, 1185-1189 (1984).
- Kaneko, H., Perez-Rojas, G., Sasasuki, T., Benike, C. J. & Engelman, E. G. *J. Immun.* **131**, 158-164 (1983).
- Kan, E. A. R., Platzter, E., Welte, K. & Wang, C. Y. *J. Immun.* **133**, 2979-2985 (1984).
- Kranz, D. M., Tonegawa, S. & Elsen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7922-7926 (1984).
- Hoffman, R. W., Bluestone, J. A., Leo, O. & Shaw, S. *J. Immun.* (in the press).
- Karpovsky, B., Titus, J. A., Stephany, D. A. & Segal, D. M. *J. exp. Med.* **160**, 1686-1701 (1984).
- Biddison, W. E., Rao, P. E., Talle, M. A., Goldstein, G. & Shaw, S. *J. exp. Med.* **159**, 783-797 (1984).
- Titus, J. A., Sharrow, S. O., Connolly, J. M. & Segal, D. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 519-523 (1981).
- Tsoukas, C. P., Carson, D. A., Fong, S. & Vaughan, J. H. *J. Immun.* **129**, 1421-1425 (1982).
- Landegren, U. *et al. J. exp. Med.* **155**, 1579-1584 (1982).
- Shaw, S., Goldstein, G., Springer, T. A. & Biddison, W. E. *J. Immun.* **134**, 3019-3026 (1985).
- Shaw, S., Johnson, A. H. & Shearer, G. M. *J. exp. Med.* **152**, 565-580 (1980).
- Sachs, D. H., Mayer, N. & Ozato, K. in *Monoclonal Antibodies and T-Cell Hybridomas* (eds Hammerling, G., Hammerling, U. & Kearny, J. F.) 95-101 (Elsevier, New York, 1981).
- Unkeles, J. C. *J. exp. Med.* **150**, 580-596 (1979).
- Segal, D. M. & Hurwitz, E. *Biochemistry* **15**, 5253-5258 (1976).
- Perez, P., Bluestone, J. A., Stephany, D. A. & Segal, D. M. *J. Immun.* **134**, 478-485 (1985).
- Staerz, U. D., Kanagawa, O. & Bevan, M. J. *Nature* **314**, 628-631 (1985).

Pre-B cells in κ-transgenic mice

Ursula Storb, Kathleen A. Denis*, Ralph L. Brinster† & Owen N. Witte*

Department of Microbiology and Immunology, University of Washington, Seattle, Washington 98195, USA

* Department of Microbiology and Molecular Biology Institute, University of California, Los Angeles, California 90024, USA

† Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Recent experiments have shown that the microinjected κ-chain gene of transgenic mice is expressed in a tissue-specific fashion only in B lymphocytes^{1,2}. The next step was to determine whether, within the B-lymphocyte lineage, the κ-chain gene was expressed in a normal developmental fashion. Normally, only μ heavy(H)-chain genes, and not κ-chain genes, are expressed in pre-B cells³. To obtain cloned cell lines derived from early cells of the B-cell lineage, we transformed bone marrow cells from κ-transgenic mice with Abelson murine leukaemia virus (A-MuLV) and tested the resultant cell lines for the retention of the κ transgene and its expression in RNA and protein. We found that cells with the pre-B phenotype exist in κ-transgenic mice. We further observed that in A-MuLV-transformed cell lines from a κ-transgenic mouse with a high copy number of the transgene, the proportion of cell lines expressing κ (transgenic κ) was higher than in cell lines from normal or low copy number transgenic mice.

Bone marrow was obtained from two 4-week-old transgenic mice and one normal sibling from the same litter. The parents were a male transgenic mouse (no. 48, ref. 2) and a normal female C57BL mouse. A portion of the bone marrow cells was directly transformed by A-MuLV⁴; the rest was first cultured for 3-4 months *in vitro*⁵ and then transformed with A-MuLV.

Kidney DNA from the three donor mice was tested for the presence of the transgene by dot hybridization with a constant region (C_κ) probe and pBR322 as described previously¹. One donor had not inherited the transgene (normal littermate), whereas the two transgenic donors had different copy numbers of the transgene (see below). Apparently the father, no. 48 (ref. 2), contained mosaic germ cells, whereas only the high copy transgenes were observed in the tail DNA (R. L. O'Brien and U.S., unpublished).

All the A-MuLV-transformed cell lines from normal and transgenic mice produce μ-chain RNA (Fig. 1, Table 1). The levels of μ RNA are variable: some cells produce very low

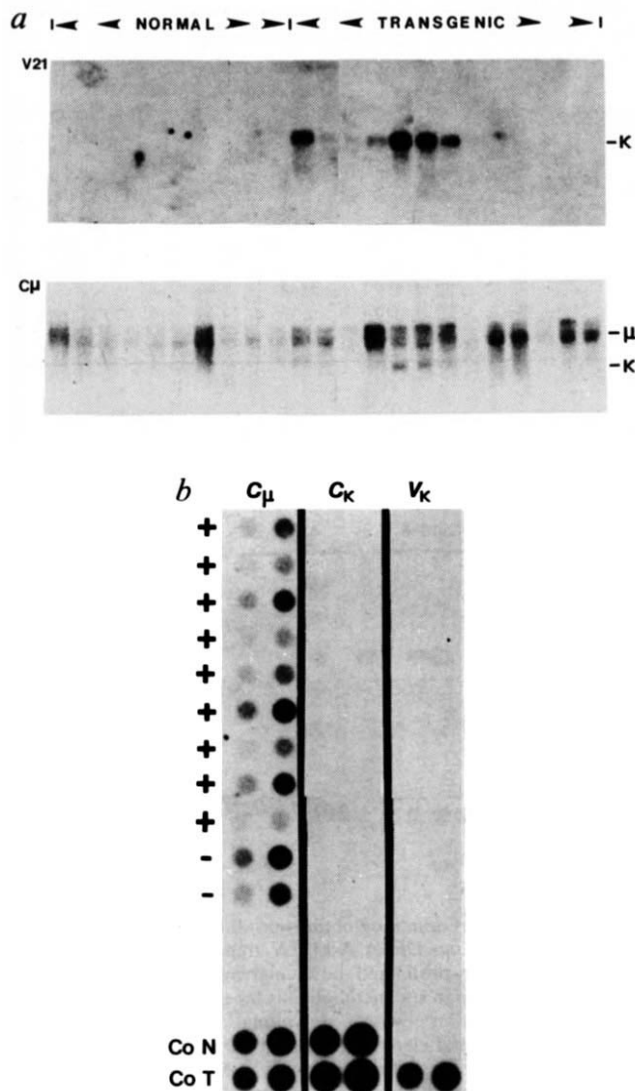


Fig. 1 RNA analysis. **a**, Northern blot of RNA from A-MuLV-transformed cell lines obtained directly from fresh bone marrow (see text). Each lane contains 10 μ g total cellular RNA. The first 10 lanes are from cell lines of normal littermates, the next 7 are from transgenic cell lines of the high copy number mouse, the last 6 are from the low copy number mouse. Top panel shows hybridization with a M.13 DNA probe containing the V_{κ} region of MOPC-21 (the transgene)¹. Bottom panel shows the same blot, first hybridized with V_{κ} M.21 and then rehybridized with a C_{μ} -containing probe⁹. Symbols μ , κ , positions of normal μ and κ mRNAs. **b**, Dot hybridization of RNA from A-MuLV-transformed cell lines. These clones were derived by A-MuLV transformation of bone marrow cells cultured for 3-4 months *in vitro* (see text). The dots were hybridized with the probes indicated: C_{μ} , C_{κ} and V_{κ} M.21; RNAs from transgenic mice (+) and normal littermates (-). CoN, control RNA from normal spleen; CoT, control RNA from transgenic spleen. Each RNA sample was dotted in two dilutions (1 and 3 μ g for C_{μ} and C_{κ} ; 2 and 6 μ g for V_{κ}).

amounts, which may in some cases be the reason for the lack of detectable μ protein (see below). The μ RNA size appears normal, probably representing the secreted and membrane forms of the protein. The cell lines also have rearranged H-chain genes on both chromosomes, indicated by hybridization with a joining segment (J_H) probe (see Fig. 2a, b for the long-term bone marrow culture-derived cell lines; data not shown for the direct transformants). Thus, the presence of the κ transgene does not prevent the rearrangement or expression of endogenous μ -chain genes, nor does it seem to impose the rearrangement of particular H-chain genes; among the different A-MuLV-transformed cell lines we found several different patterns of J_H -gene rearrange-

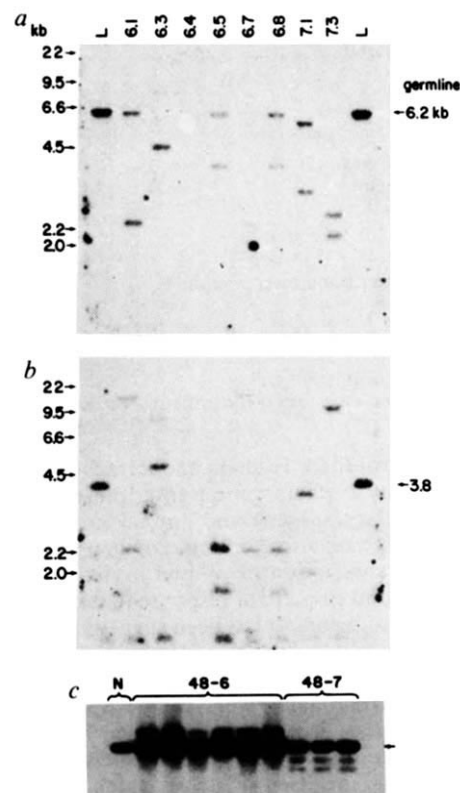


Fig. 2 Southern blots of DNAs of A-MuLV-transformed cell lines. DNAs were digested with *Eco*RI (**a**) or *Xba*I (**b**) and Southern blots were prepared as described previously¹⁰ and probed with nick-translated DNA corresponding to the J_H2 , J_H3 and J_H4 regions¹⁰. DNAs were from cell lines transformed by A-MuLV after prolonged culture of the bone marrow cells (see text). L, liver; 6.1-6.8, cell lines derived from mouse 48-6; 7.1 and 7.3, from mouse 48-6. It seems that 6.1 = 6.7 and 6.4 = 6.5 = 6.8. DNAs from direct A-MuLV transformants were digested with *Eco*RI and probed with a C_{κ} probe cloned in M13 (ref. 1). **c**, Cell lines are from normal littermates (N) and from transgenic mice 48-6 and 48-7. Arrow, tandemly repeated multiple copies of the κ transgene (14.8 kb) (~19 copies in 48-6, 5 in 48-7 cell lines) and germ line C_{κ} genes (15.2 kb).

ment. This provides further evidence that most lines are probably derived from transformation of different, independent cells, but there may be some duplication in the cell clones (Fig. 2a, b). Most of the cell lines produce μ protein (Fig. 3, Table 1), indicating that the μ RNA is functional. These results are analogous to previous findings with more mature B cells from transgenic mice that were immortalized by fusion with a hybridoma cell⁶, making it unlikely that the results described here are an artefact of the method of immortalization.

The differential copy number of the κ transgene was retained in the direct A-MuLV-transformed cell lines from the two mice (Fig. 2c). This Southern blot also shows additional bands, probably integration fragments, which differ in the cell lines from mice 48-6 and 48-7, suggesting that the transgenes are probably integrated at different sites. With respect to κ RNA, the cell lines from the two transgenic mice behaved differently; 71% of cells from the low copy number mouse 48-7, like the cells from the normal mouse, did not produce κ RNA despite the presence of κ transgenes (Table 1). The proportion of κ -producing cells from mouse 48-7 is within the range normally found in A-MuLV-transformed bone marrow cells³. Thus, even though the κ transgene is expressed in a tissue-specific manner², its presence in a B cell is not sufficient for its expression at a detectable level. True pre-B cells exist in 48-7 κ -transgenic mice. Because μ RNA is produced in these pre-B cells, transcriptional factors for H-chain genes must be present. The lack of transcription from the rearranged κ transgene suggests that different transcriptional

Table 1 Expression of κ - and μ -chain genes in A-MuLV-transformed cells from κ -transgenic mice

Cell donor	No. of cell lines*	C_{κ}	RNA V_{κ} M.21	C_{μ}	Protein† κ	μ	Pre-B (μ RNA only) total
Direct A-MuLV lines							
Normal littermate	12-	1+	0	11+ 1 (ND)	0/5+	3/5+	11/12 (92)
Transgenic no. 48-6 (high copy)	13+	12+	11+	13+	9/13+	11/13+	1/13 (8)
Transgenic no. 48-7 (low copy)	7+	2+	1+	7+	1/7+	6/7+	5/7 (71)
A-MuLV lines after bone marrow culture							
No. 48-6	7+	7-	7-	7+	0/7+	4/7+	7/7 (100)
No. 48-7	2+	2-	2-	2+	0/2+	1/2+	2/2 (100)

* -, Without; +, with transgene.

† Not all cell lines were tested for protein. ND, not done.

factors are responsible for H-chain and κ transcription and that in pre-B cells only H-chain gene transcriptional factors exist. Alternatively, the κ transgene and normal κ -chain genes may not be in an open chromatin conformation in pre-B cells. Lastly, the differential transcription of μ - and κ -chain genes may be related to differential strength of response to the same transcriptional factors, or cofactors. It has been suggested that the tissue-specific transcriptional 'enhancer' of H-chain genes is ~20 times stronger than the analogous enhancer of κ -chain genes⁷. If the amount of transcriptional factors in pre-B cells is limiting, transcription from μ -chain genes would be favoured over κ -chain genes. The amount of μ RNA in the cell lines is at least six times lower than in spleen (Fig. 1b), which indicates a lower and perhaps limiting level of transcriptional factors.

Limiting amounts of common transcriptional factors for μ - and κ -chain genes could also explain why there was a high frequency of κ RNA-producing cells in the high copy number transgenic cell lines (Fig. 1a, Table 1). Most of the direct A-MuLV cell lines from mouse 48-6 contained detectable levels of κ RNA. The cells from mouse 48-6 have about four times as many copies of the transgene as those from mouse 48-7 (Fig. 2c). Even though the κ -transcriptional enhancer may be weaker than that of μ , its presence in ~19 copies in cells of mouse 48-6 may effectively compete with μ for occasional transcription, leading to a detectable level of κ RNA.

It seems unlikely that the difference in κ frequency of cell lines derived from mice 48-6 and 48-7 results from complete transcriptional inactivation of the transgenes in the low copy number mouse. First, we found one cell line from mouse 48-7 which transcribed the κ transgene (Table 1). Second, the levels of κ RNA encoded by the transgene in spleen cells from the two mice was the same (data not shown). Besides differential transcription, differential processing or stability of the messenger RNAs could be responsible for the relative levels of μ and κ RNAs. This would require specific processing factors and/or nucleases for μ and κ RNAs, and for the κ RNAs in the κ -positive and κ -negative A-MuLV B-cell lines. These possibilities need to be explored further, but other evidence indicates an important role for transcriptional control in the regulation of immunoglobulin mRNA levels⁸.

The distribution of κ -producing cells from mouse 48-6 between the cell lines derived by the two methods of transformation was strikingly different. No κ -producing cells were found in the transformants from prolonged bone marrow cultures (Fig. 1b, Table 1). Surprisingly, these cell lines from mouse 48-6 had lost most of the copies of the transgene and only retained a low copy number, similar to the cell lines from mouse 48-7 (not shown). This finding strengthens the argument that in the high copy cell lines the large number of κ -chain genes effectively competes for limiting transcriptional factors (see above), but alternative explanations are under investigation.

In conclusion, the data reported here support the concept of sequential activation of first H-, then κ -chain genes³. The molecular basis for this control, which aids allelic exclusion of κ genes⁶, remains to be determined. To effect allelic exclusion

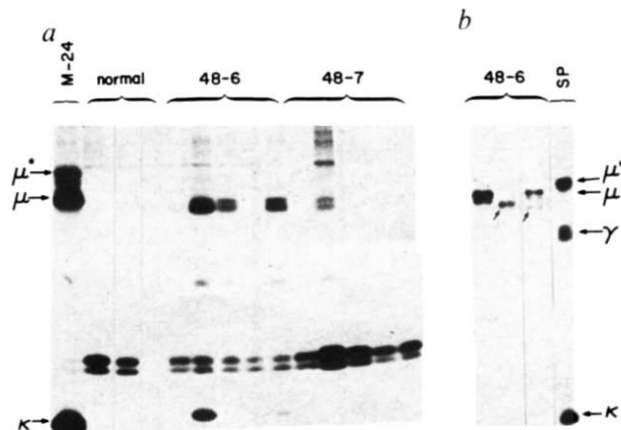


Fig. 3 Analysis of immunoglobulins produced by A-MuLV transformed cell lines. *a*, Direct A-MuLV transformed; *b*, A-MuLV transformed after prolonged bone marrow culture. Cells were grown in tunicamycin and metabolically labelled with ³H-leucine¹⁰. Immunoprecipitation with polyclonal rabbit anti-mouse immunoglobulin and electrophoresis on SDS-polyacrylamide gels was carried out as described previously¹⁰. M-24 is an established A-MuLV transformed B-cell line. Normal, 48-6 and 48-7 are cell lines derived from the three respective mice. SP, normal spleen (not treated with tunicamycin); μ^* and μ , glycosylated and unglycosylated forms of μ , respectively. Oblique arrows indicate lighter forms of μ , which were consistently observed in these two cell lines; presumably they represent translates from incomplete μ RNAs (such as *D-J-C- μ* -encoded). The double bands above the κ -chains are unexplained artefacts seen in some, but not all, immunoprecipitates with the same antiserum.

of κ genes⁶, transcriptional factors for κ -chain genes seem to be required at a high level either before or at the same time as κ -chain gene rearrangement.

We thank R. L. O'Brien for restriction analysis of DNA from high and low copy number transgenic mice and Joyce St Claire for technical assistance. This work was supported by NIH grants CA/A1 25754, CA35979, DE 02600 to U.S.; NIH grant HD 17321 and NSF grant PCM 81-07172 to R.L.B.; and grants from the March of Dimes and NIH to O.N.W., K.A.D. is a Lievre Fellow and O.N.W. is a Faculty Scholar of the American Cancer Society.

Received 14 January; accepted 1 May 1985.

- Brinster, R. *et al.* *Nature* **306**, 3332-3336 (1983).
- Storb, U., O'Brien, R., McMullen, M., Gollahan, K. & Brinster, R. *Nature* **310**, 238-241 (1984).
- Alt, F., Rosenberg, N., Lewis, S., Thomas, E. & Baltimore, D. *Cell* **27**, 381-390 (1981).
- Whitlock, C., Ziegler, S., Treiman, L., Stafford, J. & Witte, O. *Cell* **32**, 903-911 (1983).
- Whitlock, C. & Witte, O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3608-3612 (1982).
- Ritchie, K., Brinster, R. & Storb, U. *Nature* **312**, 517-520 (1984).
- Picard, D. & Schaffner, W. *Nature* **307**, 80-82 (1984).
- Nelson, K., Mather, E. & Perry, R. *Nucleic Acids Res.* **12**, 1911-1923 (1984).
- Storb, U., Arp, B. & Wilson, R. *Nature* **294**, 90-92 (1981).
- Denis, K., Treiman, L., St. Claire, J. & Witte, O. *J. exp. Med.* **160**, 1087-1101 (1984).

A transposon-like element in human DNA

K. Eric Paulson*, Niren Deka*, Carl W. Schmid*§, Ravi Misra†, Christian W. Schindler†, Mark G. Rush†, Lisa Kadyk‡ & Leslie Leinwand‡

* Department of Chemistry, University of California, Davis, California 95616, USA

† Department of Biochemistry, New York University School of Medicine, 550 First Avenue, New York City, New York 10016, USA

‡ Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

§ To whom correspondence should be addressed.

Mobile genetic elements have been reported in prokaryotes, plants, yeast and *Drosophila* (reviewed in refs 1, 2). The only transposon-like sequences reported for mammalian organisms are closely related to retroviruses³⁻⁷, although undoubtedly other transposon families exist within the mammalian genome. Although mobile genetic elements can only be identified as such if their mobility can be demonstrated in existing populations, transposon and transposon-like elements share several common biochemical and structural features^{1,2}. Here we demonstrate that a repetitive human sequence has many of the diagnostic features of transposable elements. This 2.3-kilobase (kb) transposon-like element contains two flanking long terminal repeat (LTR)-like 350-base pair (bp) repetitive sequences, each of which begins with the sequence 5' TG... and ends with ...CA 3'. The transposon-like element is bounded by 5-bp direct repeats. Discrete-length polyadenylated transcripts from HeLa cells are homologous to the transposon-like element. Members of this transposon-like family are found in extrachromosomal circular DNA molecules.

Sun *et al.*⁸ previously described the 'O' family of repetitive sequences. A slightly different representation of two members of this family (see below) indicates that the 350-bp O family repeats begin with TG and end with CA (B. Baum, personal communication), a hallmark of the LTRs that flank transposable

elements¹⁻³. Therefore, the O repeats may be solitary LTRs, and additional O family members could be linked to longer transposon-like elements. To test this interpretation, we determined the DNA base sequence of additional members of the O family, which were isolated as described previously⁸.

One such recombinant DNA clone, λ THE 1-A, contains two members of O at opposite ends of a 1.9-kilobase (kb) *Bgl*II restriction fragment (Fig. 1a). (This structure is called a transposon-like human element, and corresponding clones are designated THE 1; the O family is now called the 'O-LTR' family.) Subcloned restriction fragments of λ THE 1-A (Fig. 1a) were used as hybridization probes for additional λ THE clones as well as genomic DNA fragments. A 2-kb sequence flanked by two O family repeats is also observed in clones λ THE 1-B and -C (Fig. 1a). Because these three recombinant DNAs share homologous sequences and similar maps, they must be members of a repetitive sequence family. The following hybridization experiments confirm the repetitive nature of the THE 1 family.

Human DNA digested with the restriction enzymes listed in Fig. 1a was hybridized to O-LTR repeats and subcloned restriction fragments of λ THE 1-A (Fig. 1a). The *Bgl*II digest of human DNA illustrates the main results of our study (Fig. 1b). The 350-bp O-LTR repeat hybridizes to a highly repetitive 1.9-kb *Bgl*II fragment in addition to restriction fragments of other lengths (Fig. 1b), indicating that O-LTR repeats are associated with longer repetitive sequences. Subclones derived from the longer repetitive element of λ THE 1-A hybridize to many of the same genomic restriction fragments as the O-LTR family, with noticeably lower background (Fig. 1a, b). Comparison of the λ genomic clones THE 1-A, -B and -C reveals restriction site polymorphisms (Fig. 1a), many of which correspond to minor genomic restriction fragments that can be arranged as a consensus restriction map (Fig. 1a). Evidently, most of the sequences that are homologous to THE 1 family repeats share a single consensus sequence.

Dot-blot and library screening hybridizations indicate that there are ~30,000 O-LTR repeats and 10,000 copies of the internal element within human DNA (data not shown). These

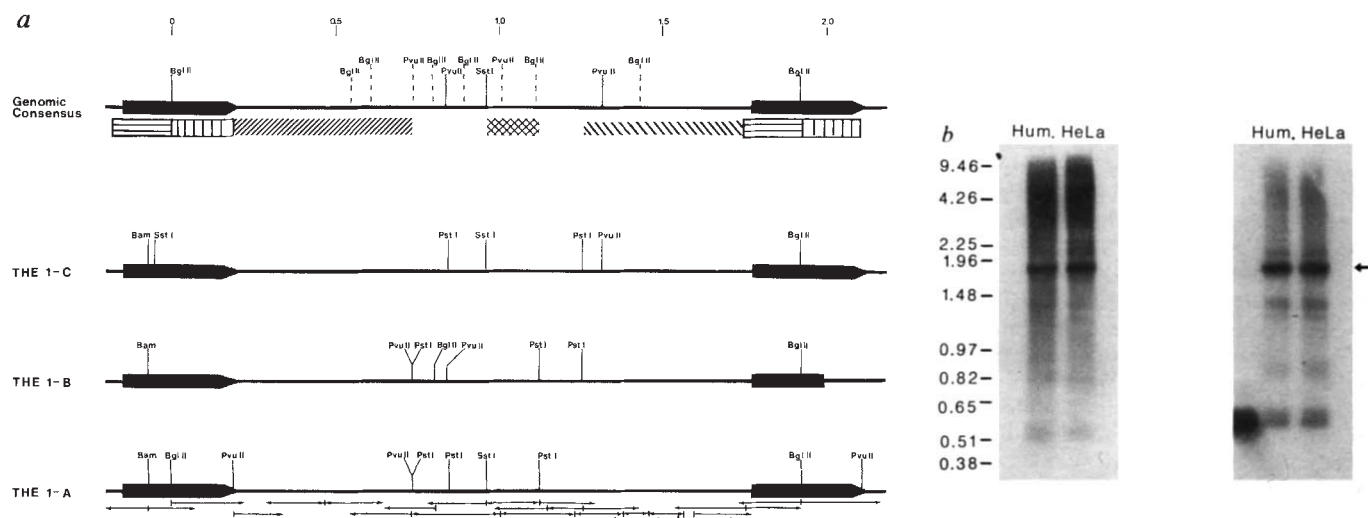


Fig. 1 *a*, Genomic and clone restriction maps. Restriction fragments of THE 1-A (indicated schematically by the cross-hatched and shaded bars under the genomic consensus map) were subcloned and used as hybridization probes to the indicated restriction digests of λ THE 1-A, -B and -C and genomic DNA. Genomic digests are complicated by the appearance of a number of weak and presumably polymorphic restriction fragments (*b*). Relatively rare sites are indicated in the consensus map by dotted lines; strong sites are indicated by solid lines. The right LTR of λ THE 1-B is deleted, a finding which has been confirmed by limited base sequencing. The base sequencing strategy used for λ THE 1-A is indicated by the arrows positioned under the map of λ THE 1-A. These restriction fragments were subcloned into M13 and their sequences were determined by Messing's modifications of the Sanger dideoxynucleotide procedure and limited Gilbert-Maxam sequencing of refractory regions. *b*, Mapping of genomic restriction digests. Placental and cell culture DNAs (labelled human and HeLa respectively) digested with *Bgl*II were hybridized to the LTR family (pBRO₄ in ref. 8) and pUC19 subclone of the 1.9-kb *Bgl*II fragment of λ THE 1-A as depicted in *a*. A highly repetitive 1.9-kb *Bgl*II fragment, as indicated by the arrow, as well as several minor bands are revealed by this hybridization. The autoradiograph was exposed overnight without an intensifying screen. The minor bands can be rationalized according to the rarer sites in the consensus map of *a*.

1 2

—

—28S

—18S

Fig. 3 Representation of THE 1 family in HeLa cell transcripts. pUCO₄ DNA was radiolabelled by nick-translation and hybridized to poly(A)⁻ (lane 1) and poly(A)⁺ (lane 2) from HeLa cells. The pUCO₄ subclone contains the entire transposon and one composite LTR as explained in the text. 28S and 18S correspond to ribosomal RNA migration. Cytoplasmic RNA was isolated from HeLa cells and fractionated by oligo(dT)-cellulose chromatography into poly(A)⁻ and poly(A)⁺ fractions. Aliquots (5 µg) of each fraction were fractionated on a 1.0% formaldehyde agarose gel in MOPS buffer and transferred to nitrocellulose. Hybridization was carried out in 5 × SSC, 0.01 M NaPO₄ pH 7.4, 2 × Denhardt's, 100 µg ml⁻¹ denatured salmon sperm DNA, 10% dextran sulphate at 60 °C. Washes were carried out in 2 × SSC. 0.2% SDS at 60 °C.

Blot hybridization has been used to test whether THE 1 family members, in analogy to other families of transposons, code for transcription products^{2,3,9,10}. A discrete 2.0-kb transcript in poly(A)⁺ cytoplasmic RNA from HeLa cells hybridizes to a subclone derived from λ THE 1-A (Fig. 3). This transcript is absent from poly(A)⁻ cytoplasmic RNA. The subclone used in this hybridization, pUCO₄ (a 1.9-kb *Bgl*II fragment of λ THE 1-A in the plasmid pUC) contains the entire transposon-like sequence in addition to one composite copy of O-LTR (see map in Fig. 1a). Perhaps because they used only the O-LTR sequence as a hybridization probe, Sun *et al.*⁸ did not observe this tran-

The structure of the THE 1 family of repetitive sequences is distinct from that of other families of repeats of retroviruses within the human genome. Abundant families of human repeats such as the *Alu* family or *LINE-1* family do not have flanking LTRs and share other structural features that differ from those of THE 1 repeats^{8,14,15}. The marked similarities between retroviral proviruses and transposable elements demonstrate their close relationship^{2,3}. Certainly, many transposon-like structures observed in mammalian DNAs are variants of retroviral proviruses⁴⁻⁷. However, the base sequence of THE 1 repeats does not exhibit homology to any known mammalian retrovirus, and the sequence length of this family is significantly shorter than

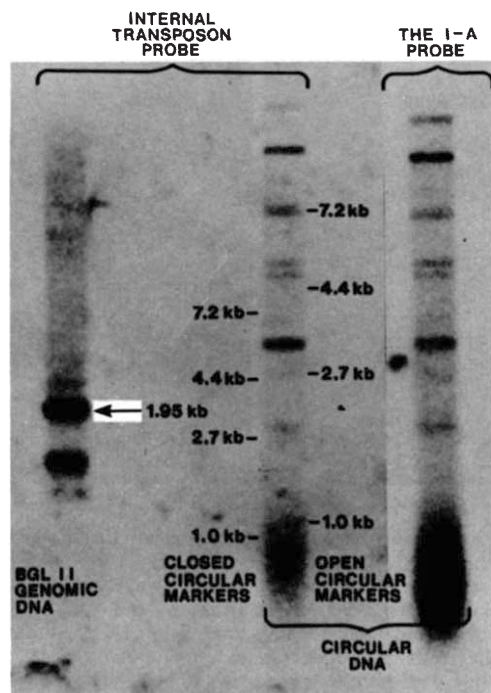


Fig. 4 Covalently closed circular DNAs from the Jurkat lymphoblastoid human cell line were isolated and resolved by gel electrophoresis^{12,13}. Blots of these gels were hybridized to subclones corresponding to an internal 0.38-kb *Pst*I fragment of λ THE 1-B (left panel as indicated), the 1.9-kb *Bgl*II fragment of λ THE 1-A (right panel as indicated) or the LTR-specific hybridization probe (pBR0₄ of ref. 8; data not shown). The right and left panels, which correspond to gels that were run for slightly different times, are juxtapositioned to indicate the coincidence of the bands. Circular DNA marker lengths are indicated in the left panel, which also includes a *Bgl*II digest of genomic DNA revealing the diagnostic 1.9-kb band described in Fig. 1a. As an important negative control, neither pBR nor M13 DNAs exhibited any hybridization to these gels.

that of typical retroviruses. Therefore, THE 1 family repeats should be regarded as a new class of repetitive transposon-like sequences.

Blot hybridization reveals the THE 1 structure in monkey but not in rodent DNA, so that the mobility of these sequences over long evolutionary times seems certain. However, blot hybridization of a sequence flanking THE 1-B to the DNAs of 19 individuals indicates that this particular member of the family is fixed in the human genome (data not shown). The mobility or relative immobility of this family remains to be determined.

This research was supported by USPHS grants GM21346 (to C.W.S.) and GM29090 (to L.L.), and by NSF grant PCM-8203006 (to M.G.R.). Ch.W.S. and R.M. are trainees supported by USPHS NIH training grants 5T32GM07308 and 5T32GM07238-08.

Received 11 March; accepted 22 May 1985.

- Flavell, A. J. & Ish-Horowitz, D. *Nature* **292**, 591-595 (1981).
- Temin, H. M. *Cell* **21**, 599-600; 27, 1-3 (1980).
- Varmus, H. E. *Science* **216**, 812-820 (1982).
- Schmidt, M., Glogler, K., Wirth, C. & Horak, I. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6696-6700 (1984).
- Strecker, R. E. *Nature* **298**, 767-769 (1982).
- Kuff, E. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1992-1996 (1983).
- Mager, D. L. & Henthorn, P. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7510-7514 (1984).
- Sun, L., Paulson, K. E., Schmid, C. W., Kadyk, L. & Leinwand, L. *Nucleic Acids Res.* **12**, 2669-2690 (1984).
- Elder, R. T., Loh, E. Y. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2432-2436 (1983).
- Flavell, A. J., Levis, R., Simon, M. A. & Rubin, G. M. *Nucleic Acids Res.* **9**, 6279-6291 (1981).
- Rose, A. M. & Snutch, T. P. *Nature* **311**, 485-486 (1984).
- Krolewski, J. J. & Rush, M. G. *J. molec. Biol.* **174**, 31-40 (1984).
- Schindler, C. W. & Rush, M. G. *J. molec. Biol.* (in the press).
- Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065-1070 (1982).
- Singer, M. F., Thayer, R. E., Grimaldi, G., Leeman, M. I. & Fanning, T. G. *Nucleic Acids Res.* **11**, 5739-5745 (1983).

Lymphokine dependence of *in vivo* expression of MHC class II antigens by endothelium

G. Groenewegen, W. A. Buurman & C. J. van der Linden

Department of Surgery, Biomedical Center, Hospital St Annadal, University of Limburg, Maastricht, The Netherlands

Changes in the expression of class II antigens of the major histocompatibility complex (MHC) have an integral role in the regulation of immune responses, and are brought about *in vitro* by soluble mediators¹⁻⁴. However, the mechanism that underlies *in vivo* expression of MHC class II antigens in, for example, endothelial cells in the absence of immunological stimulation has not been studied. We demonstrate here that expression of MHC class II antigens is not a constitutive property of endothelial cells, for MHC class II antigen-positive endothelial cells do not express these antigens during treatment with the immunosuppressive agent cyclosporin A. *In vivo* MHC class II antigen expression by canine endothelial cells is therefore dependent on factors, probably the lymphokine γ -interferon produced by the immune system, whose secretion is inhibited by cyclosporin A.

γ -Interferon (IFN- γ) has been identified as the lymphokine responsible for the augmented MHC class II antigen expression observed on monocytes *in vitro*⁴, although recently an intermediate messenger, produced by murine tumour macrophages in response to IFN- γ , has been reported to be involved in this process⁵. The MHC class II antigen-inducing capacity of IFN- γ has also been clearly demonstrated *in vitro* for endothelial cells from human umbilical vein⁶ and fibroblasts⁷. Furthermore, there is evidence that the expression of these antigens changes on immunological stimulation⁸⁻¹². For example, rat skin^{9,10} and gut epithelium¹⁰, both normally MHC class II antigen-negative cells, express these antigens during graft-versus-host disease and local infections¹¹. We therefore investigated whether the normal 'physiological' expression of MHC class II antigens is dependent on factors such as IFN- γ .

An *in vivo* model was selected that assessed the effect of lymphokines on normal endothelial cell MHC class II antigen expression by analysing the response to cyclosporin A; this compound has been shown to inhibit the *in vitro* production of the factor responsible for MHC class II antigen expression¹³⁻¹⁵. The canine model was chosen because the dog resembles humans¹⁶⁻¹⁹ in exhibiting MHC class II antigens on the capillary endothelial cells of the skin and the glomerular and peritubular capillary endothelium of the kidney.

MHC class II antigen expression was studied in skin and kidney biopsies from five dogs treated with cyclosporin A. The drug was given once daily as an oral solution (Sandoz), and drug blood levels were monitored by HPLC and maintained between 0.2 and 0.4 mg ml⁻¹. The cyclosporin A doses administered ranged from 15.0 to 22.5 mg per kg per day. Drug treatment did not affect the health of the dogs. Renal function, as assessed by serum creatinine and urea concentrations, was unaffected; creatinine concentrations were 40-60 μ mol l⁻¹ during treatment (normal range in dogs 30-120 μ mol l⁻¹). Wound healing of the biopsy incisions was also unimpaired.

The expression of MHC class II antigens during cyclosporin A treatment was studied in tissue biopsies by indirect immunofluorescence using two monoclonal anti-MHC class II antibodies (Fig. 1). Renal biopsies taken from the first dog studied on days 10, 19 and 32 of treatment revealed that MHC class II antigens were present before treatment (Fig. 1a), but could not be detected during treatment (Fig. 1b); the antigens were consistently absent from the renal vasculature in all biopsies taken after 10 days of cyclosporin A treatment.

Endothelial cells were identified using the marker protein, FVIII-related antigen²⁰ (Fig. 2). In renal biopsies from untreated dogs, both antibodies stained endothelium of glomerular capil-

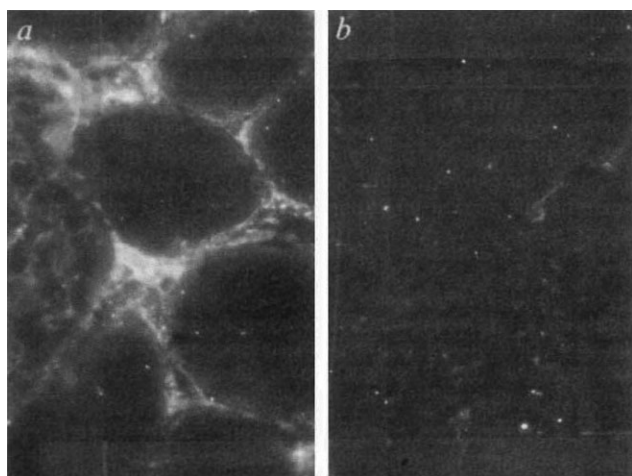


Fig. 1 Distribution of MHC class II antigens in the kidney before (a) and during (b) treatment with cyclosporin A ($\times 400$). Monoclonal antibodies 7.2 (ref. 25) and 7.5.10.1 (ref. 26), known to cross-react with canine MHC class II antigens²⁷, were used in the ascites form. Both monoclonal antibodies gave identical patterns of immunofluorescence. The reactivity of antibody 7.5.10.1 is shown. a, A biopsy taken before treatment, showing part of a glomerulus (left margin) and several unstained tubules; the glomerular endothelium stained for MHC class II antigens, as did peritubular capillary endothelium. Biopsy b, taken after 10 days of cyclosporin A administration, is devoid of antigens reacting with anti-MHC class II reagents, even at antibody concentrations 10-fold higher than the concentration that appeared on titration to be optimal in biopsies of untreated dogs.

Methods. Healthy beagles from the colony of the University of Maastricht were regularly screened for infections; the last vaccination had been given at least 12 months before the experiment. Tissue specimens, obtained by open biopsies, were snap-frozen in isopentane, cooled on dry ice. Tissue sections ($4\ \mu\text{m}$) were fixed in acetone at -20°C and rehydrated in phosphate-buffered saline (PBS) at room temperature. Subsequently, sections were incubated for 30 min at room temperature with the desired antibody, washed for 15 min in PBS and incubated with an appropriate fluorescein-labelled second antibody. Antibodies were used at twice the concentration shown by titration to be still saturating, to achieve maximal sensitivity of the fluorescence technique. After a final wash the section was embedded in glycerol.

laries and vasculature around the tubules, whereas tubular cells and basement membrane structures were not stained. Before and during treatment with cyclosporin A, the distribution of FVIII-related antigen in endothelium was unchanged. However, MHC class II antigen distribution differed before and during treatment. Our findings were confirmed in a second dog treated with cyclosporin A and biopsied on days 10, 20 and 33 of treatment. In three other dogs, MHC class II antigens could not be detected in renal biopsies taken on day 20 of treatment, whereas the distribution of FVIII-related antigen remained unchanged.

Table 1 summarizes the results of skin and kidney biopsies. The expression of FVIII-related antigen was similar before and during treatment. However, endothelial expression of MHC class II antigens was absent in biopsies taken on day 10 of treatment, whereas before treatment the antigens were clearly present on endothelial cells. Continued treatment until day 25 resulted in a continued absence of the antigens on endothelial cells, whereas the distribution of FVIII-related antigen was unaffected. The fact that the distribution of this marker antigen was unchanged with treatment makes it unlikely that the disappearance of MHC class II antigens from the vasculature of both skin and kidney was caused by a toxic effect of cyclosporin A on endothelium.

In contrast to the disappearance of the endothelial MHC class II antigens, the expression of these antigens by a second vascular bed, the Langerhans cells of the skin, was unaffected by cyclo-

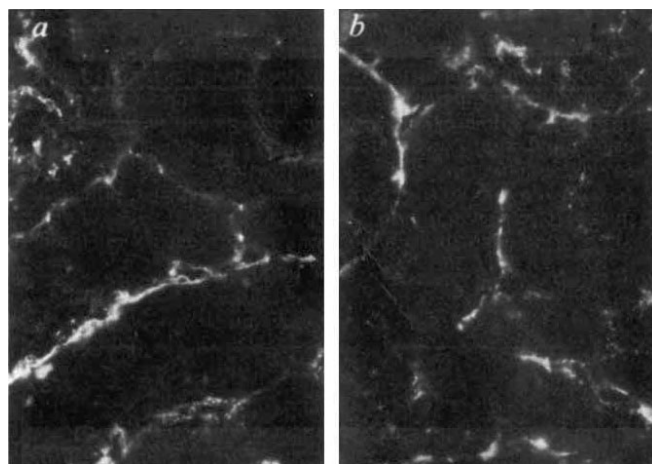


Fig. 2 Distribution of FVIII-related antigen in the kidney before (a) and during (b) treatment with cyclosporin A ($\times 400$). Antiserum specific for dog FVIII-related antigen²⁸ was used to identify endothelium in tissue sections. In biopsies a and b, taken, respectively, before and after 10 days' cyclosporin A administration, endothelial cells of glomerular and peritubular capillaries are stained. In a, a glomerulus is partially visible in the upper left-hand corner; b, glomeruli are visible in the upper right-hand corner and on the left margin of the photomicrograph.

sporin A treatment (Table 1). Whereas treatment for 10 days resulted in the loss of antigens from endothelium, treatment for as long as 25 days caused no change in MHC class II antigen expression on skin Langerhans cells.

In vitro MHC class II antigen expression on endothelial cells is induced and maintained by lymphokines. Cyclosporin A does not influence lymphokine-induced MHC class II antigen expression on endothelial cells *in vitro*, but has an effect via inhibition of the production of lymphokines responsible for antigen expression^{13,14}. Similarly, the expression of MHC class II antigens on endothelium *in vivo* seems to be lymphokine-dependent, as treatment with cyclosporin A, which inhibits the production of the relevant factors, results in the observed decreased expression of MHC class II antigens on endothelium in skin and kidney. Previous studies have already suggested that the immune system affects the expression of these antigens⁸⁻¹². The present study now shows that the immune system sustains the normal expression of MHC class II antigens on endothelium by means of lymphokines.

Although the nature of the lymphokine(s) involved *in vivo* is unknown, it seems likely to be IFN- γ , because this lymphokine is known to induce the expression *in vitro* of MHC class II antigens^{4,6,7} and its production is inhibited by cyclosporin A¹⁵. Our data strongly suggest that basal secretion of IFN- γ ²¹ is the cause of the 'normal' MHC class II antigen expression on endothelial cells. The expression of MHC class II antigens *in vivo* on capillary and venous endothelium and not on arterial endothelium suggests a difference in the sensitivity of these cells to IFN- γ , since, *in vitro*, antigen expression by arterial endothelium can be induced by lymphokines^{13,14}. During immunological stimulation, high concentrations of IFN- γ can induce the expression of MHC class II antigens on cells with low sensitivity to this lymphokine. Differences in reactivity to lymphokines might be part of the mechanism regulating the expression of MHC class II antigens on different cell types and therefore immunological responsiveness.

Our experiments demonstrate that regulation of MHC class II antigen expression on endothelial cells and on Langerhans cells of the skin is different, the former being abrogated by cyclosporin A, with the latter being unaffected. The consequence of this difference in the allograft situation will depend on the species. In the rat model it can be assumed that the donor passenger cells will be replaced by recipient cells. Because only

Table 1 Expression of MHC class II antigens in skin and kidney biopsies from dogs, before and during treatment with cyclosporin A

	Kidney		Skin	
	Glomerular endothelium	Peritubular endothelium	Capillary endothelium	Langerhans cells
Before treatment	+	+	+	+
During treatment	-	-	-	+

MHC class II antigen expression was studied with monoclonal antibodies using an indirect immunofluorescence technique. The presence or absence of antigens is indicated, respectively, by + and -.

passenger cells express MHC class II antigens in the rat kidney, the disappearance of these cells results in complete loss of donor-type MHC class II antigens, leading to long graft survival²². In man and dog, however, the expression of MHC class II antigens by endothelial cells prevents the disappearance of donor-type MHC class II antigens from an allograft. The present data show that removal can be achieved by cyclosporin A, thus suggesting an important new mode of action for this reagent in prolonging allograft survival.

We believe this to be the first demonstration that the expression of MHC class II antigens *in vivo* on endothelial cells is not a constitutive property of these cells. The expression *in vivo* on capillary endothelium in the canine kidney and skin is sustained by the immune system via lymphokines. The presence of MHC class II antigens on the capillary endothelial cells may be a consequence of the rheological characteristics of the vessels lined by these cells. The conditions for interaction between lymphocytes and endothelial cells, which are known to function as antigen-presenting cells^{23,24}, are most favourable on these MHC class II-positive capillary endothelial cells.

The variations in the expression of MHC class II antigens by endothelial cells as a result of products of the immune system, suggest new insights into the control of immune responses, and offer interesting opportunities for evaluating immune status in patients. Moreover, the abrogation of endothelial MHC class II antigen expression caused by cyclosporin A offers a new explanation for its immunosuppressive effect.

We thank Mrs G. M. A. A. Jeunhomme, Mrs A. J. J. M. Daemen and Mrs E. E. M. Spronken for technical assistance and Mrs K. Spronck for secretarial assistance. Cyclosporin A was a gift of Mr Stokvis (Sandoz). Cyclosporin A concentrations were determined by the Pharmacological and Toxicological Laboratory, Hospital St Annadal, Maastricht, The Netherlands. Antiserum against canine FVIII-related antigen was a gift of Professor Dr Bouma. This work was supported by grant C81.297 from Nier Stichting Nederland and grant 29.57.10 from Fungo.

Received 17 December 1984; accepted 3 May 1985.

- Steinman, R. M., Nogueira, N., Witmer, M. D., Tydings, J. D. & Melman, J. S. *J. exp. Med.* **152**, 1248-1261 (1980).
- Scher, M. G., Beller, D. J. & Unanue, E. R. *J. exp. Med.* **152**, 1684-1698 (1980).
- Stegg, P. S., Moor, R. N. & Oppenheim, J. J. *J. exp. Med.* **152**, 1734-1744 (1980).
- Stegg, P. S., Moor, R. N., Johnson, H. M. & Oppenheim, J. J. *J. exp. Med.* **156**, 1780-1793 (1982).
- Walker, E. B., Maino, V., Sanchez-Lanier, M., Warner, N. & Stewart, C. J. *exp. Med.* **159**, 1532-1547 (1984).
- Pober, J. S. *et al. J. exp. Med.* **157**, 1339-1353 (1983).
- Pober, J. S. *et al. Nature* **305**, 726-729 (1983).
- Wadgymar, A., Urmsion, J., Bauml, R. & Halloran, P. F. *J. Immun.* **132**, 1826-1832 (1984).
- Lampert, J. A., Suitters, A. J. & Chisholm, P. M. *Nature* **293**, 149-150 (1981).
- Mason, D. W., Dallman, M. & Barclay, A. N. *Nature* **293**, 150-151 (1981).
- Barclay, A. N. & Mason, D. W. *J. exp. Med.* **156**, 1665-1676 (1982).
- de Waal, R. M. W. *et al. Nature* **303**, 426-429 (1983).
- Groenewegen, G., Buurman, W. A. & Jeunhomme, G. M. A. *Transplant Proc.* **17**, 1417-1418 (1985).
- Groenewegen, G., Buurman, W. A., Jeunhomme, G. M. A. & van der Linden, C. J. *Transplantation* (in the press).
- Reem, G. H., Cook, L. A. & Vilček, J. *Science* **221**, 63-65 (1983).
- Natali, P. G., De Martimo, C., Marcellini, M., Quaranta, V. & Ferrone, S. *Clin. Immun. Immunopath.* **20**, 11-20 (1981).
- Paul, L. C., van Es, L. A. & Baldwin, W. A. III, *Clin. Immun. Immunopath.* **19**, 206-223 (1981).
- Scott, H., Brandtzaeg, P., Hirschberg, H., Solheim, B. G. & Thorsby, E. *Tissue Antigens* **18**, 195-202 (1981).
- Fuggle, S. V. *et al. Transplantation* **35**, 385-390 (1983).
- Hoyer, L. W., de los Santos, R. P. & Hoyes, J. R. *J. clin. Invest.* **52**, 2737-2744 (1973).
- Stewart, W. E. II, in *The Interferon System*, 55-57, 75-76 (Springer, Berlin, 1979).
- Fabre, J. W. *Transplantation* **34**, 223-236 (1982).

- Hirschberg, H., Bergh, O. J. & Thorsby, E. *J. exp. Med.* **152**, 249S-255S (1980).
- Groenewegen, G. & Buurman, W. A. *Scand. J. Immun.* **19**, 269-273 (1984).
- Deeg, H. J. *et al. Immunogenetics* **16**, 445-457 (1982).
- Koning, F., Schreuder, J., Giphart, M. & Bruning, H. *Hum. Immun.* **9**, 221-230 (1984).
- Doveren, R. F. C., Buurman, W. A., Schutte, B., Groenewegen, G. & v. d. Linden, C. J. *Tissue Antigens* **25**, 255-265 (1985).
- Bouma, B. N., Dodds, W. J., van Mourik, J. A., Sixma, J. J. & Webster, W. P. *Scand. J. Haemat.* **17**, 263-275 (1976).

Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II

Jeffrey Bonadio & Peter H. Byers

Departments of Pathology and Medicine and the Center for Inherited Disease, University of Washington, Seattle, Washington 98195, USA

Although the perinatal lethal form¹ of osteogenesis imperfecta (OI type II) occasionally results from large rearrangements within the genes encoding type I collagen²⁻⁵, most mutations are far more subtle. The complexity of the human collagen genes⁶ precludes cloning and sequencing each gene from every patient, and we have therefore developed an approach to localizing mutations at the protein level. We report here that cells cultured from 15 infants with OI type II synthesized both normal type I procollagen and a form that was unstable, poorly secreted and excessively modified. Abnormal procollagen from different strains was overmodified to different extents. The patterns of overmodification we observed are best explained by mutations that disrupt the Gly-X-Y sequence of pro α chains, and thus alter the rate of propagation of triple helix from COOH-terminus to NH₂-terminus⁷. As a consequence, a given mutation allows overmodification of all three chains in a molecule NH₂-terminal to its position in the triple helix.

After pulse labelling with ³H-proline for 16 h, each of the 15 OI cell strains secreted pro α 1(I) and pro α 2(I) collagen chains with normal mobility and others with delayed mobility (Fig. 1a). The OI cells retained more pro α 1(I) and pro α 2(I) chains than control cells, and most of the intracellular procollagen consisted of chains with delayed mobility.

After treatment with pepsin, which removes the NH₂- and COOH-terminal propeptide extensions from type I procollagen molecules in triple helical conformation⁸, the OI α 1(I) and α 2(I) chains migrated as doublets (Fig. 1b). One chain from each doublet co-migrated with α 1(I) and α 2(I) chains from control cells and the other chain migrated more slowly than normal.

In the presence of α , α' -dipyridyl, which inhibits most of the intracellular modification of type I procollagen⁹, the OI cells synthesized only pro α 1(I) and pro α 2(I) chains, which co-migrated with control chains (Fig. 1c). Thus, the slow mobility of the abnormal pro α 1(I) and pro α 2(I) chains was apparently due to excessive modification, and the mutation in each OI cell strain described here was not a large insertion (or deletion^{10-14,29}) but was either a single amino-acid substitution or an insertion or deletion that we could not detect by SDS-polyacrylamide vertical slab gel electrophoresis (PAGE)¹⁵.

To determine the distribution of overmodification along the triple helix, we cleaved collagen-sized molecules asymmetrically with fibroblast collagenase (Fig. 1d). For some cell strains (for convenience these strains are designated group C), the abnormal α 1(I) and α 2(I) chains yielded a COOH-terminal B fragment and NH₂-terminal A fragment¹⁶, which were both overmodified. For other cell strains, only the NH₂-terminal A fragment was overmodified (groups B and A).

We further defined the extent of increased modification by the relative migration of each of four detectable CNBr peptides from abnormal α 1(I) chains (Fig. 1). For group C cell strains the entire triple helix was overmodified, for group B cell strains overmodification began near the fibroblast collagenase cleavage site, and for group A cell strains overmodification began

Fig. 1 a, Electrophoretic separation of pro α chains synthesized by OI and control cells. Three patient groups were identified based on the distribution of excessive post-translational modification within the triple helical domain (see Figs 2, 4). These groups designate biochemically distinguishable cell strains and do not correspond to the recently proposed clinical sub-classification of OI type II²⁵. Indeed, all patients described here fit the OI type IIA classification. There were seven cell strains in group A, three in group B and five in group C. The data from one representative of each group are shown. Two points emerge from these data. First, the electrophoretic mobilities of the pro α 1(I) and pro α 2(I) chains of some molecules are not normal. Second, there is considerable intracellular retention of abnormal pro α 1(I) and pro α 2(I) chains. The identity of bands with normal and delayed mobility was determined by CNBr peptide mapping (data not shown). **b**, Electrophoretic separation of α chains. The pro α chains synthesized by the OI type II cells were converted into two forms of α 1(I) and α 2(I), with one form of each chain migrating more slowly than normal and the other form co-migrating with the control (arrows indicate medium and cell-layer α 1(I) chains with delayed mobility). The delayed mobility of the abnormal α 1(I) and α 2(I) chains differed among the three groups; the delay was greatest for the abnormal α chains synthesized by cells from group C, intermediate for α chains from group B and least for α chains from group A. Resistance to peptic attack indicated that the abnormal pro α chains were incorporated into molecules that were triple helical in the conditions of digestion.

c, Electrophoretic separation of unmodified pro α chains. Only pro α 1(I) and pro α 2(I) chains that co-migrated with control chains were synthesized by the OI cell strains. The extent of hydroxylation of prolyl and lysyl residues and glycosylation of hydroxylysyl residues of type I procollagen synthesized by OI type II cells in culture has been measured directly by Steinmann *et al.*¹⁸ and by Bateman *et al.*²¹. Both studies found that prolyl hydroxylation was slightly increased over control levels and that lysyl hydroxylation and hydroxylysyl glycosylation were markedly increased, the overmodified pro α chains migrated more slowly on SDS-PAGE than their normal counterparts, and the difference in mobilities was abolished if the OI cells were pulse-labelled in the presence of α , α' -dipyridyl. **d**, Electrophoretic separation of α -chain fragments after cleavage with fibroblast collagenase. Type I collagen molecules synthesized by cell strains from group C yield A- and B-chain fragments that migrate as doublets (arrow indicates overmodified α 2(I)B fragment). By contrast, only the A fragments from collagen synthesized by cell strains from groups B and A migrate as doublets while the B fragments co-migrate with the control.

Methods. **a**, Dermal fibroblast cell strains were established from explants of skin as described previously^{10,15}. Cells were plated at a density of 2.5×10^5 in 35-mm culture dishes and allowed to attach and spread overnight. The cells were preincubated for 2–4 h in 0.7 ml of serum-free Dulbecco's modified Eagle's medium (DMEM) containing $50 \mu\text{g ml}^{-1}$ of ascorbic acid. The medium with ascorbic acid was replaced with fresh medium and ascorbate and $[2,3,4,5\text{-}^3\text{H}]$ proline (101 Ci mmol^{-1} ; Amersham) was added for 17 h. The medium and cell layer were then collected separately into protease inhibitors¹⁵, dialysed into 1 mM ammonium bicarbonate containing the same inhibitors, and lyophilized. The dried samples were dissolved in electrophoresis sample buffer, and the proteins were denatured by boiling for 5 min. The proteins were separated in reducing conditions by SDS-PAGE with 2 M urea added to the buffers to enhance separation of pro α chains²⁶. Gels were processed for fluorography with ENHANCE (NEN). **b**, Medium and cell-layer proteins were collected as described above and the proteins were denatured by boiling for 5 min. The proteins were separated in reducing conditions by SDS-PAGE with 2 M urea added to the buffers to enhance separation of pro α chains²⁶. Gels were processed for fluorography with ENHANCE (NEN). **c**, Cells in culture were plated as before and then preincubated for 2 h in 0.7 ml of serum-free DMEM and 0.5 mM α , α' -dipyridyl (Sigma). The cells were then pulse-labelled with ^3H -proline in fresh medium plus α , α' -dipyridyl for 4 h. The cell-layer proteins were collected as described above and the constituent proteins were separated by SDS-PAGE in reducing conditions. Pro α chains from the medium of control cells labelled for 4 h in the absence of α , α' -dipyridyl have been separated in an adjoining lane for comparison. **d**, Collagen-sized molecules from the medium (after pepsin digestion) were dialysed into 50 mM Tris-HCl buffer pH 7.5, containing 10 mM CaCl_2 , then digested for 24 h at room temperature with fibroblast collagenase (final concentration $\sim 1.0 \mu\text{g}$ collagenase per $10 \mu\text{g}$ collagen). Differences in the relative migration of A- and B-chain fragments produced by the control and OI cell strains result from slight differences in the composition of SDS-PAGE.

NH₂-terminal to this cleavage site within the domain of the triple helix defined by the α 1(I)CB7 peptide.

The increase in post-translational modification was accompanied by a decrease in thermal stability of the abnormal collagen molecules (Fig. 3). Molecules secreted by cells from group C had melting temperatures in the range 36–38°C, those secreted by cells from group B melted in the 38–39°C range, and those secreted by cells from group A melted in the 39–41°C range. The melting temperature of normal collagens secreted by OI cell strains and collagens secreted by control cell strains was 41.5–42.5°C. Thus, the abnormal OI molecules with the lowest melting temperatures were overmodified to the greatest extent.

These OI type II cell strains in culture synthesize unstable procollagen molecules that are excessively modified and poorly secreted. Because only some of the type I procollagen molecules from each strain are affected and because overmodification is

often asymmetrically distributed within the triple helical domain, we consider that the mutations harboured by these strains are in the genes encoding the pro α chains themselves. Their principal effect is to destabilize the triple helix of molecules incorporating a mutant chain.

Formation of a collagen triple helix requires a glycine in every third position¹⁷, and the structure is stabilized by hydrogen bonds that depend on hydroxyprolyl residues in the Y-position of repeating Gly-X-Y triplets. Hydroxylation of prolyl and lysyl residues begins on nascent chains and ceases when a stable triple helical conformation is achieved. Mutations which result in substitution for a glycine or in small deletions or insertions that disrupt the Gly-X-Y sequence could each produce an unstable triple helical domain and, as a consequence, cause the intracellular retention and excessive modification of affected molecules. As overmodification is often asymmetrically dis-

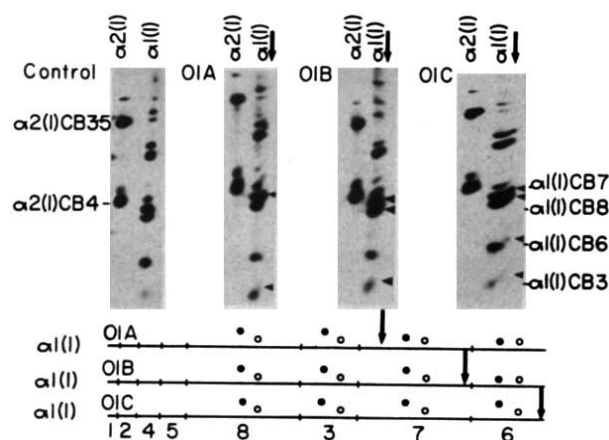


Fig. 2 CNBr peptides of $\alpha 1(I)$ and $\alpha 2(I)$ chains. We describe the migration of CNBr peptides derived only from $\alpha 1(I)$ chains because this chain yields four peptides from the region of the type I procollagen molecule that we are studying, whereas $\alpha 2(I)$ chains yield only two large peptides. The overmodified $\alpha 1(I)$ chain from group C OI cell strains yielded an array of CNBr peptides similar to the control, except that each peptide was delayed in mobility. The overmodified $\alpha 1(I)$ chain from group B cell strains yielded $\alpha 1(I)$ CB7, $\alpha 1(I)$ CB8 and $\alpha 1(I)$ CB3 peptides with delayed mobility and an $\alpha 1(I)$ CB6 peptide which co-migrated with $\alpha 1(I)$ CB6 from the normal $\alpha 1(I)$ chain synthesized by the OI cells and from the control $\alpha 1(I)$ chain. The overmodified $\alpha 1(I)$ chain from group A cell strains yielded a single peptide in the region of $\alpha 1(I)$ CB7 and $\alpha 1(I)$ CB8 that migrated slightly more slowly than the normal $\alpha 1(I)$ CB7 peptide, a peptide which migrated more slowly than $\alpha 1(I)$ CB3, and a peptide that co-migrated with $\alpha 1(I)$ CB6. We have shown by CNBr peptide mapping of these α chains after cleavage with fibroblast collagenase (see also ref. 15) that the large peptide spot migrating slightly more slowly than the normal $\alpha 1(I)$ CB7 peptide consisted of overmodified $\alpha 1(I)$ CB7 and $\alpha 1(I)$ CB8 peptides (data not shown). Taken together with the results of fibroblast collagenase cleavage, we concluded that the abnormal type I procollagen molecules synthesized by the various OI cell strains were not overmodified to the same extent; for group C the entire triple helix was overmodified, for group B overmodification began near the fibroblast collagenase cleavage site, and for group A overmodification began NH_2 -terminal to this site within the domain of the triple helix defined by the $\alpha 1(I)$ CB7 peptide. The line diagrams of the $\alpha 1(I)$ chains indicate the COOH-terminal end of the domain of increased post-translational modification (arrows). The vertical relationship between closed (OI) and open (normal) circles indicates the relative migration of CNBr peptides during electrophoresis.

Methods. Proteins were separated in the first dimension on 5% SDS-PAGE in denaturing but nonreducing conditions. The lanes were cut out and the proteins were digested in the gel with CNBr (100 mg ml^{-1} in 70% formic acid; Sigma) for 2 h at room temperature^{10,15,27}. Each gel strip was washed with water for three 20-min periods, equilibrated for 20 min with sample buffer lacking SDS and urea, then placed horizontally over a 12.5% gel (second dimension) that contained no urea. The strips were overlaid with sample buffer containing dithiothreitol and 5% SDS, and the CNBr peptides were separated by electrophoresis.

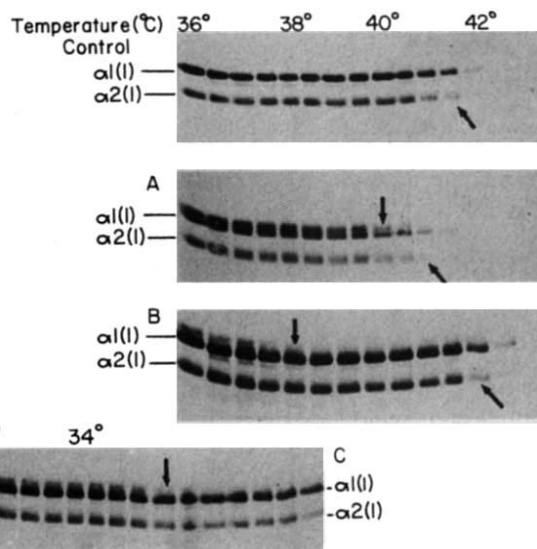


Fig. 3 The thermal denaturation temperature (T_m) of collagens secreted by representatives of OI type II cells in groups A, B and C.

Methods. Pepsin-digested procollagens from the medium were dissolved in $700 \mu\text{l}$ of 0.4 M NaCl , 0.01 M Tris-HCl , $\text{pH } 7.4$. The samples were placed in a water bath at 30°C and then warmed to 43°C at a rate of 5°C h^{-1} . Aliquots ($36 \mu\text{l}$) were removed at 0.5°C increments, rapidly cooled to 20°C and then digested with trypsin ($100 \mu\text{g ml}^{-1}$ final concentration; Worthington) for 2 min at 22°C ²⁸. The reaction was stopped by adding phenylmethanesulphonyl fluoride and SDS to achieve a final concentration of 2 mM and 2%, respectively. The whole mixture was placed in boiling water for 3 min. The proteins were separated by electrophoresis on 5% SDS-polyacrylamide gels containing 0.5 M urea . Fluorograms were quantitated by scanning gel densitometry. The slanted arrows below the top three fluorograms indicate the T_m of normal collagen molecules; the vertical arrows indicate the T_m of each abnormal molecule. In a separate experiment (data not shown) we have shown that the T_m of normal collagen secreted by the representative cell strain from group C was 41.5°C .

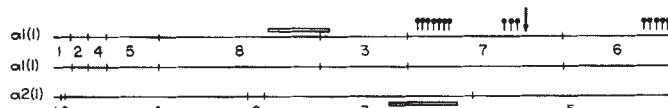


Fig. 4 Nature and location of mutations in the α chains of type I collagen synthesized by the cells from 17 patients with type II osteogenesis imperfecta. Circles indicate the sites at which increased post-translational modification begins in type I collagen molecules from 15 OI type II cell strains for which the mutation does not change the apparent length of pro α chains. Although arrayed along the $\alpha 1(I)$ chain, our data do not distinguish between mutations in $\alpha 1(I)$ or $\alpha 2(I)$ chains. The arrow indicates the fibroblast collagenase cleavage site. Bars indicate the approximate size and location of the deletion from $\alpha 1(I)$ and $\alpha 2(I)$, respectively, in two OI type II cell strains discussed previously^{10,11,22}.

tributed, we think that triple helix formation, which normally begins at the COOH-terminus of the helix and propagates toward the NH_2 -terminus⁷, proceeds at an appropriate rate until the region of the mutation is reached. Thereafter, the rate of formation of a stable triple helix is delayed and (because both pro $\alpha 1(I)$ and pro $\alpha 2(I)$ chains are excessively modified) all chains in abnormal molecules are available to the post-translational modifying enzymes for longer than normal^{15,18-21}. Assuming heterozygosity^{15,18} and the random assortment of nascent pro α chains in the rough endoplasmic reticulum as pro $\alpha 1(I)$ and pro $\alpha 2(I)$ chains associated to form trimers, only a portion of the procollagen molecules synthesized would be expected to contain a mutant chain.

These concepts are supported by studies which showed that a cysteine substitution within the triple helix of $\alpha 1(I)$ chains resulted in the overmodification, intracellular retention and instability of all molecules that contained the mutant chain¹⁸. Sequence analysis (D. Cohn, B. Steinmann, R. E. Gelinas and P. H. Byers, manuscript in preparation) has identified a single base substitution in a glycine codon that results in a Gly $\rightarrow$ Cys substitution 27 amino acids from the COOH-terminus of $\alpha 1(I)$ chains encoded by this allele. As the entire length of the triple helix of abnormal molecules was shown to be overmodified, the

position of the mutation in this cell strain is correlated precisely with the extent of overmodification. It is also notable that the abnormal type I procollagen molecules synthesized by OI cell strains harbouring large deletions are also overmodified, and the extent of overmodification can be correlated with the position of the mutation in the molecule.

Mutations that do not change the apparent length of the pro α chains of type I procollagen account for most of the mutations which give rise to the OI type II phenotype (Fig. 4). To date, cell strains from only three infants with OI type II have been shown to synthesize procollagen molecules that harbour deletions from either pro $\alpha 1(I)$ or pro $\alpha 2(I)$ ^{10-14,22,29}. In contrast, the 15 cell strains described here, 8 others described in detail elsewhere^{15,18,21} and 43 others that we have partially characterized (unpublished) synthesized molecules in which the mutation does

not result in a detectable change in pro α chain length. While the nature and location of these mutations will have to be determined by gene or messenger RNA sequence studies, mapping the sites of structural alterations in mutant proteins augments S₁ nuclease mapping strategies^{23,24} and identifies a small region within a large collagen gene for detailed study.

We thank the many physicians and families who provided us with skin biopsies from infants with OI type II, Dr Eugene Bauer for providing the fibroblast collagenase, and Nancy Higgins and William Dickie for preparing the manuscript. This work was supported in part by grants from the NIH (AM-21557, AM-07171, and GM-15253) and by a Clinical Research Grant (6-298) from the March of Dimes Birth Defects Foundation. J.B. is the recipient of an NIH individual fellowship award (AM-07171); P.H.B. was an Established Investigator of the American Heart Association during the tenure of this study.

Received 14 December 1984; accepted 14 May 1985.

1. Silience, D. O., Senn, A. S. & Danks, D. M. *J. med. Genet.* **16**, 101-116 (1979).
2. Hollister, D. W., Byers, P. H. & Holbrook, K. A. *Adv. hum. Genet.* **12**, 1-83 (1982).
3. Byers, P. H. & Bonadio, J. F. in *International Medical Reviews, Pediatrics 5, Genetic and Metabolic Disease* (eds Scriver, C. R. & Lloyd, J.) (Butterworths, London, in the press).
4. Prockop, D. J. & Kivirikko, K. *New Engl. J. Med.* **311**, 376-386 (1984).
5. Prockop, D. J. *J. clin. Invest.* **75**, 783-787 (1985).
6. Tate, V., Finer, M., Boedtker, H. & Doty, P. *Cold Spring Harb. Symp. quant. Biol.* **47**, 1039-1049 (1982).
7. Bachinger, H. P., Fessler, L. I., Timpl, R. & Fessler, J. H. *J. biol. Chem.* **256**, 13193-13199 (1981).
8. Bornstein, P., Kang, A. H. & Piez, K. A. *Biochemistry* **5**, 3803-3812 (1966).
9. Kivirikko, K. I. & Myllylä, R. *Meth. Enzym.* **82**, 245-304 (1982).
10. Barsh, G. S. & Byers, P. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5142-5146 (1981).
11. Barsh, G. S., Roush, C. L., Bonadio, J., Byers, P. H. & Gelinas, R. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2870-2874 (1985).
12. Williams, C. J. & Prockop, D. J. *J. biol. Chem.* **258**, 5915-5921 (1983).
13. Chu, M.-L. *et al. Nature* **304**, 78-80 (1983).
14. de Wet, W. J., Pihlajaniemi, T., Myers, J., Kelly, T. E. & Prockop, D. J. *J. biol. Chem.* **258**, 7721-7728 (1983).
15. Bonadio, J., Holbrook, K. A., Gelinas, R. E., Jacob, J. & Byers, P. H. *J. biol. Chem.* **260**, 1734-1742 (1985).
16. Stricklin, G. P., Eisen, F. Z., Bauer, E. A. & Jeffrey, J. J. *Biochemistry* **17**, 2331-2337 (1978).
17. Prockop, D. J., Kivirikko, K. I., Tuderman, L. & Guzman, N. A. *New Engl. J. Med.* **301**, 13-23 (1979).
18. Steinmann, B. *et al. J. biol. Chem.* **259**, 11129-11138 (1984).
19. Trelstad, R. L., Rubin, D. & Gross, J. *Lab. Invest.* **36**, 501-508 (1977).
20. Kirsch, E. *et al. Eur. J. clin. Invest.* **11**, 39-47 (1981).
21. Bateman, J. F., Mascara, T., Chan, D. & Cole, W. G. *Biochem. J.* **217**, 103-115 (1984).
22. Byers, P. H. *et al. Am. J. hum. Genet.* **35**, 39A (1983).
23. Dixon, L. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4525-4528 (1984).
24. Pihlajaniemi, T. *et al. J. biol. Chem.* **259**, 12941-12944 (1984).
25. Silience, D. O., Barlow, K. K., Garber, A. P., Hall, J. G. & Rimoin, D. *Am. J. med. Genet.* **17**, 407-423 (1984).
26. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
27. Barsh, G. S., Peterson, K. E. & Byers, P. H. *Collagen and Related Research* **1**, 543-548 (1981).
28. Bruckner, P., Eikenberry, E. F. & Prockop, D. J. *Eur. J. Biochem.* **118**, 607-613 (1981).
29. Chu, M.-L., Gargiulo, Y., Williams, C. & Ramirez, F. *J. biol. Chem.* **260**, 691-694 (1985).

Sliding distance of actin filament induced by a myosin crossbridge during one ATP hydrolysis cycle

Toshio Yanagida, Toshiaki Arata & Fumio Oosawa

Department of Biophysical Engineering and Science, Osaka University, Toyonaka, Osaka, Japan

Muscle contraction results from a sliding movement of actin filaments induced by myosin crossbridges on hydrolysis of ATP^{1,2}, and many non-muscle cells are thought to move using a similar mechanism³⁻⁵. The molecular mechanism of muscle contraction, however, is not completely understood^{6,7}. One of the major problems is the mechanochemical coupling at high velocity under near-zero load⁸⁻¹³. Here, we report measurements of the sliding distance of an actin filament induced by a myosin crossbridge during one ATP hydrolysis cycle in an unloaded condition. We used single sarcomeres from which the Z-lines, structures which anchor the thin filaments in the sarcomere, had been completely removed by calcium-activated neutral protease (CANP)¹⁴ and trypsin, and measured both the sliding velocity of single actin filaments along myosin filaments and the ATPase activity during sliding. Our results show that the average sliding distance of the actin filament is ≥ 600 Å during one ATP cycle, much longer than the length of power stroke of myosin crossbridges deduced from mechanical studies of muscle, which is of the order of 80 Å (for example, ref. 15).

Myofibrils were prepared from glycerinated crab (*Portunus trituberculatus*) leg muscle. Because the crab muscle has a long sarcomere (~ 7 μ m; thin and thick filaments are 2.7 and 4.8 μ m long, respectively), the sarcomere shortening takes a long time and the shortening velocity can be measured using an optical microscope. Figure 1a, b shows phase contrast and fluorescence images, respectively, of a myofibril labelled with phalloidin tetramethylrhodamine (PDTMR). In Fig. 1b, thin filaments only are bright because PDTMR specifically binds F-actin¹⁶. Figure 1c is a fluorescence micrograph of single sarcomeres in myofibrils from which the Z-lines have been removed. When 5 mM ATP was added to the single sarcomeres in the absence of Ca²⁺ at 0°C, thin filaments were dissociated from the thick filaments and dispersed into the solution. Single F-actin filaments labelled with PDTMR can be clearly observed in solution using the fluorescence microscope¹⁷. We confirmed that the thin filaments were completely dissociated from the A-band and the dispersed filaments single, not bundled (Fig. 1d). This point was also confirmed by electron microscopy. The treatment with CANP

and trypsin thus completely digested the Z-lines, leaving no structures to interconnect the thin filaments.

Using single sarcomeres without Z-lines, we next measured the sliding velocity of single thin filaments along thick filaments and the ATPase activity during sliding. Figure 2 shows sequential fluorescence micrographs of a single sarcomere without Z-lines during shortening at 5°C. The shortening velocity of PDTMR-labelled single sarcomeres was measured directly. Shortening was initiated by rapidly adding a small drop of ATP solution (final concentration 5 mM) to single sarcomeres which had sedimented and weakly adhered to the surface of a glass cell in rigor solution containing an ATP regenerating system and 0.2 mM CaCl₂ (Fig. 2). Thin filaments immediately started to slide into the H-band and met those from the opposite side at the centre of the H-band (M-line) 0.20-0.25 s after ATP addition. The velocity decreased but the shortening continued until the filaments had overlapped. The sliding velocity (V_s) of thin filaments along thick filaments in the H-band was 4.7 ± 0.5 μ m s⁻¹ at 5°C. The shortening velocity of single sarcomeres was also measured by monitoring the increase in light-scattering intensity of the sarcomere suspension after rapidly mixing it with ATP using a stopped-flow apparatus (Fig. 3). The shortening velocity obtained from the initial rise in the light-scattering intensity (5.3 ± 0.5 μ m per half-sarcomere per s at 5°C) agreed well with that measured directly under a fluorescence microscope (Table 1). The shortening velocity of the sarcomeres was about 1.7 times higher than the maximum shortening velocity of a muscle fibre under the same conditions (3.0 ± 0.2 μ m per half-sarcomere per s at 5°C). This does not result from the removal of the Z-lines because the shortening velocity of myofibrils before the removal was 1.5-2 times faster than this. Arata *et al.* have reported that the shortening velocity of rabbit psoas myofibrils (>10 μ m per half-sarcomere per s at 0°C) is also much higher than the maximum velocity of muscle fibres (<2 μ m per half-sarcomere per s)¹⁸. These high values do not result from the fact that shortening was initiated by adding ATP to actomyosin rigor complexes in myofibrils, because the shortening velocity was the same when initiated by adding Ca²⁺ to relaxed myofibrils. Therefore, these high velocities probably result from the removal of some connective materials^{19,20} during preparation of the myofibrils. Such a high velocity is also seen in proteoplasmic streaming in *Nitella* cells³ and movement of small organelles along actin cables in *Chara* cells²¹ (60-80 μ m s⁻¹ at room temperature) which are believed to be driven by the actin-myosin sliding movement.

Figure 3 shows the time course of ADP liberation from single sarcomeres without Z-lines after they had been rapidly mixed with ATP at 5°C under the same conditions as those used for

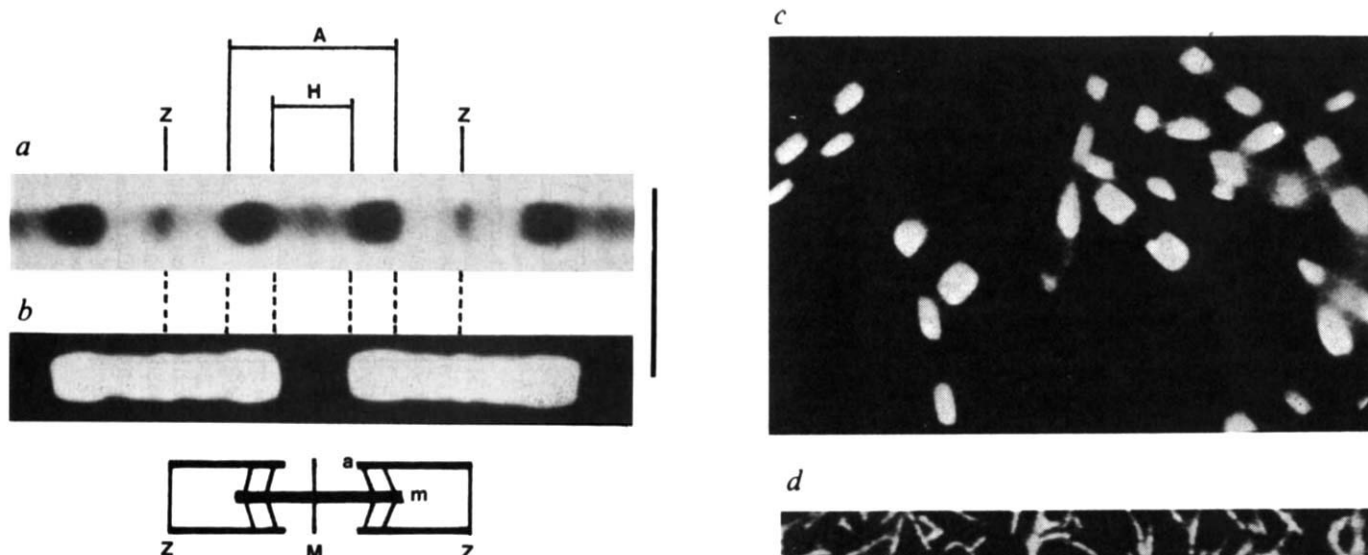


Fig. 1 *a, b*, Phase contrast and fluorescence images, respectively, of a crab myofibril labelled with PDTMR. A diagram of the sarcomere is also shown. A, A-band; H, H-band; Z, Z-line; a, thin filament; m, thick filament; M, M-line. *c*, Fluorescence images of single sarcomeres labelled with PDTMR after removal of Z-lines from myofibrils by CANP and trypsin. Note that the length of thin filaments in the single sarcomere is 2–2.5 μm , 0.2–0.5 μm shorter than the original length of 2.7 μm because part of the thin filaments at the free end (Z-line side) was mechanically broken during the preparation. Therefore, the Z-lines of the sarcomeres used were not only chemically but also mechanically removed. *d*, Fluorescence images of single thin filaments dissociated from sarcomeres in a relaxing solution (rigor (100 mM KCl, 7 mM MgCl_2 , 20 mM Tris-HCl pH 7.5) + ATP + EGTA). Note that all fluorescent filaments have the same brightness which is uniform along the filament. The filaments could not be distinguished from the fluorescence images of single thin filaments purified from rabbit skeletal muscle. Scale bars, 5 μm .

Methods. Myofibrils were prepared from glycerinated crab muscle and homogenized for 5 min in rigor solution with a Waring blender. CANP was obtained from the crab muscle and purified by the method of Ishiura *et al.*³². CANP specifically digests the Z-lines and hardly affects other protein^{14,33}. Thin filaments (F-actin) in myofibrils or sarcomeres were specifically labelled with PDTMR by immersing them in rigor solution containing $\sim 1 \mu\text{M}$ PDTMR for a few minutes at 0 °C. Treatment with PDTMR do not affect ATPase activity in the presence of Ca^{2+} (ref. 34). Fluorescent myofibrils and sarcomeres were observed with an inverted microscope (Zeiss IM 35) equipped with epifluorescence optics, a Zeiss Neofluor $\times 100$ objective (oil immersion, NA, 1.3), a 100 W mercury arc lamp and a Zeiss rhodamine filter ($\lambda_{\text{excitation}} = 540 \text{ nm}$; $\lambda_{\text{emission}} = 580 \text{ nm}$). Fluorescent single thin filaments were observed with a laser fluorescence microscope system described elsewhere³⁵. The fluorescent images were recorded on videotape with a high-sensitivity television camera (Ikegami SIT camera CTC-900; time resolution 17 ms) and a video recorder (Victor CR 6650L).

Table 1 Sliding distance of thin filament induced by a myosin crossbridge during one ATP cycle

Temperature (°C)	Sliding velocity of thin filament, V_s ($\mu\text{m s}^{-1}$)		ATPase activity, V_{ATP} (ADP mol per myosin mol s^{-1})*	Sliding distance of thin filament per ATP cycle, M (Å)†	
	Light scattering	Fluorescence microscope		Light scattering	Fluorescence microscope
5	$5.3 \pm 0.7(5)$	$4.7 \pm 0.8(20)$	$1.0 \pm 0.3(4)$	680 ± 240	600 ± 230
10	$12 \pm 1.5(5)$	—	—	—	—
15	$20 \pm 2.0(5)$	—	$4.1 \pm 0.5(3)$	650 ± 170	—

All observed values (V_s and V_{ATP}) are shown as mean $\pm$ half-range. Numbers in parentheses indicate the number of measurements. Values calculated from our measurements (M) are shown as the mean $\pm$ maximum possible error ($\delta M = |(\partial M / \partial V_s) \delta V_s| + |(\partial M / \partial V_{\text{ATP}}) \delta V_{\text{ATP}}|$).

* ATPase activity of each myosin molecule during sarcomere shortening was obtained as $V_{\text{ATP}} = V_s^s \times (W/C) \times 2L_0 / (L_0 + L)$, where V_s^s is ATPase activity of sarcomere shown as mol ADP per g sarcomere per s; C , myosin content (by weight) of a sarcomere; W , relative molecular mass of myosin (4.8×10^5 , ref. 28); L , length of the overlap region between thin and thick (crossbridges) filaments per half-sarcomere before shortening; L_0 , the length at full overlap of thick and thin filaments. ATPase activity in the non-overlap region (in the relaxed state) was neglected because it was expected to be >20 times smaller than that in the overlap region during shortening. The average values of L_0 and L of sarcomeres used here were 2.3 and 1.1 μm , respectively. C was calculated to be 0.35 by comparing the area of the myosin heavy chain of the sarcomere with that of purified myosin, the weight of which was found according to Arata *et al.*²⁹, on SDS gels.

† The sliding distance of thin filaments during one ATP cycle (lower limit) was obtained according to the equation shown in the text from the sliding velocity (V_s) and the ATPase activity (V_{ATP}). The average number of myosin molecules in half a thick filament overlapping with thin filaments during sliding, N_m , was obtained as $N_m = N_m \times (L_0 + L) / 2L_0$ ($= 640 \times (2.3 + 1.1) / 4.6 = 470$). The number of myosin molecules in half a thick filament, N_m , was obtained as $N_m = (L_0 / P_m) \times n$ where P_m is the myosin subunit repeat (14.3 nm; Namba *et al.*²²); and n , number of myosin molecules per repeat ($= 4$; Wakabayashi^{30,33}). The value of n was also determined to be ~ 4 (3.8–4.3) by SDS polyacrylamide gel electrophoresis according to Tregear and Squire³¹.

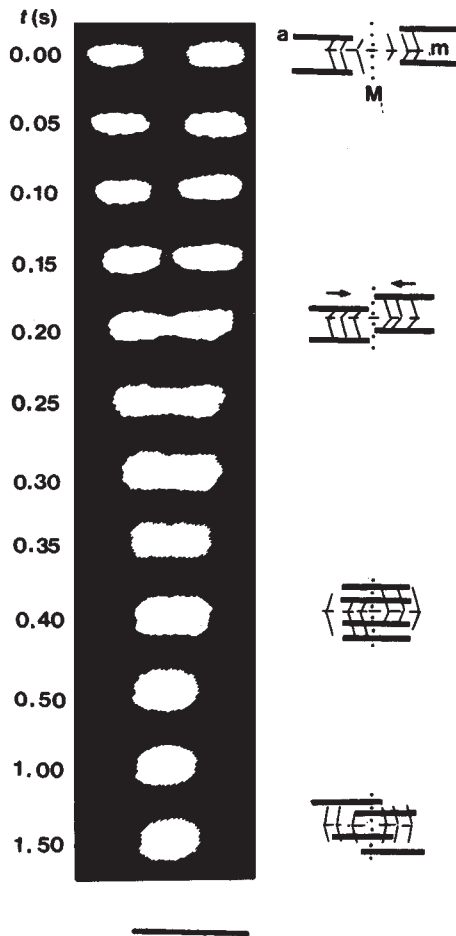


Fig. 2 Sequential fluorescence micrographs of a PDTMR-labelled sarcomere without Z-lines during shortening at 5 °C. Times (t) are shown in s. Scale bar, 5 μm . Shortening was initiated by rapidly adding ATP solution from a small pipette at $t=0$ s to the suspension of sarcomeres on a glass slide enclosed with a bank of glue. The bathing solution (rigor +0.2 mM CaCl_2) contained 0.2% β -mercaptoethanol to suppress photobleaching. At $t \geq 0.5$ s only a small spot, probably the overlap region, could be seen at the set level of brightness; the image of the ends of filaments became too vague to be visualized. Sample temperature was controlled by attaching a metal rod connected to a bath cooler to the stage of the inverted microscope.

the measurement of shortening velocity. The ATPase activity during shortening at 5 °C was 1.0 ± 0.3 mol ADP per mol myosin s^{-1} at 2 mM ATP and decreased to 0.13 ± 0.02 mol ADP per mol myosin s^{-1} after completion of the shortening.

The average sliding distance of a thin filament induced by hydrolysis of 1 mol ATP was calculated from the sliding velocity of the thin filament and the ATPase activity during sliding. The average number of myosin molecules in half a thick filament overlapped with the thin filaments during the shortening, $\bar{N}_m$, is 470 (see Table 1). The ATPase activity during the sliding, V_{ATP} , is 1 s^{-1} at 5 °C (Table 1) and the ratio of thin to thick filaments, r , is 6 (ref. 22). Therefore, the number of ATP cycles on one thin filament per s, N_{ATP} , was calculated as $N_{\text{ATP}} = (\bar{N}_m \times V_{\text{ATP}})/r = (470 \times 1 \text{ s}^{-1})/6 = 78 \text{ s}^{-1}$. The sliding distance of the thin filament during one ATP cycle, M , is calculated as $M = V_s/N_{\text{ATP}} = 4.7$ (or 5.3) $\mu\text{m s}^{-1}/78 \text{ s}^{-1} = 600$ (or 680) Å assuming that the ATP cycles occur in turn without overlapping. This value should be the lower limit because the ATP cycles sometimes occur simultaneously, in which case M would become larger. When the temperature was increased to 15 °C, a similar high value of M was obtained (Table 1). Thus, the sliding distance of the thin filament induced by the myosin crossbridge during one ATP hydrolysis cycle in the unloaded

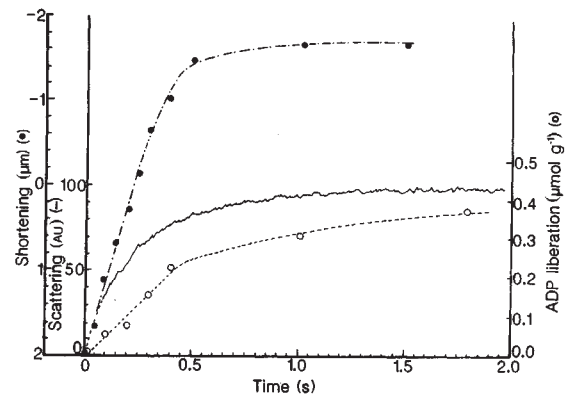


Fig. 3 Time course of shortening, light-scattering intensity and ADP liberation during contraction of single sarcomeres without Z-lines.

Methods. Closed circles, extent of shortening observed directly by an inverted fluorescence microscope (see Fig. 2). Ordinate, width between the tops of thin filaments on both sides of the sarcomere (H-band). After these filaments had overlapped, their width was measured between their tail ends and shown as negative values. Solid line, intensity of light scattering from a suspension of single sarcomeres at 90° and $\lambda = 460$ nm. The light-scattering intensity increased following an exponential equation consisting of two time constants and then decreased by >10 times slower than the fastest phase (not shown). Assuming that the fastest component of the increase in the light-scattering intensity arose from an increase in the overlap between thin and thick filaments at the A-band and the second from the overlap between the thin filaments from the two opposite sides, the sliding velocity of thin filaments was calculated as $V = (\frac{1}{2}l)/\tau_1$ ($= \frac{1}{2} \times 2.3 (\pm 0.2) \mu\text{m}/0.22 \text{ s} = 5.2 (\pm 0.5) \mu\text{m s}^{-1}$, where τ_1 is the fastest time constant and l the width of the H-band. l was measured with a fluorescence microscope after labelling with PDTMR. The suspension of sarcomeres was rapidly mixed with ATP and the time course of change in light intensity measured with a stopped-flow spectrophotometer (Union Giken RA-401). Open circles, ADP liberation, measured by the method of Arata *et al.*¹⁸. The ATPase reaction was initiated by rapidly adding a small amount of concentrated ATP solution to the 20 mg ml^{-1} suspension of sarcomeres in rigor solution containing 0.2 mM CaCl_2 , vigorously stirred and stopped by adding HCl (final concentration 0.3 M). ATP and HCl solutions were added using a small syringe connected to an electric solenoid. The ATPase reaction of the sarcomeres (width of H-band = $2.3 \pm 0.2 \mu\text{m}$) was coupled with 15 mg ml^{-1} pyruvate kinase and 2 mM phosphoenolpyruvate to regenerate ATP. The ATPase activity was determined from the rate of pyruvate liberation, measured by the method of Reynarn *et al.*³⁶. The maximum activity of pyruvate kinase (Sigma) was 50 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ in rigor solution at 5 °C. When the ATPase activity of sarcomeres of 20 mg ml^{-1} is 20 $\mu\text{M s}^{-1}$ (1.4 per mol ADP per mol myosin per s, see Table 1), the lag between the ATP hydrolysis and ATP regeneration reactions in the presence of 15 mg pyruvate kinase per ml is estimated to be only 0.015 s, taking into account that the Michaelis constant of ADP for pyruvate kinase is 0.2 mM³⁶. The protein concentration of the sarcomere was measured by the biuret method and the absorbance at 280 nm was determined after treatment with 1 M NaOH for 24 h at 20 °C (ref. 37). All experiments were performed at 5 °C.

condition is much longer than the size of the power stroke of the crossbridge, 80 Å deduced from studies on isometric tension transients of muscle (ref. 15).

Because the Reynolds number (inertial force/viscous force) is extremely small for the movement of the thin filament ($\sim 10^{-5}$) or the relaxation time of the velocity of the thin filament after removal of the external force is <1 ns ($= m/\zeta = Dm/kT = (10^{-8} \text{ cm}^2 \text{ s}^{-1} \times 10^{-16} \text{ g})/(1.38 \times 10^{-23} \text{ J K}^{-1} \times 300 \text{ K}) = 10^{-10} \text{ s}$, where m is the mass of the thin filament; ζ , the friction constant; D , translational diffusion constant; k , Boltzmann constant; and T , temperature)²³, the movement of the thin filament induced by the myosin crossbridge decays immediately (<1 ns) after its detachment and the thin filament cannot be moved by an inertial force. The possibility of the movement of several other filaments

hydrodynamically induced by the movement of a thin filament is expected to be very low for two reasons; first, the myosin heads and filament backbone that occupy most of the space in the sarcomere^{22,24} are expected to almost completely abolish the flow of the solution; second, even if the flow of solution is induced, it would almost be abolished because the directions of the flow from two sides of the sarcomere oppose each other. It is also unlikely that the high sliding velocity of the thin filaments in the sarcomere results from electrostatic interactions between them, as increasing the concentration of free Mg^{2+} from 1 to 7 mM did not increase the shortening velocity, although the attractive force among the thin filaments should be stronger at higher concentration of divalent cations²⁵. In fact the shortening velocity decreased by 10% at 100 mM KCl and 30% at 50 mM KCl. Furthermore, the fact that the elastic modulus of thin filaments obtained from analysis of their thermal bending motion in a myosin-free ghost fibre²⁶ is almost the same as that when they are independent in solution¹⁷ strongly suggests that the electrostatic interaction among thin filaments in the sarcomere is very weak.

How can such a long sliding distance of the thin filament per one ATP hydrolysis cycle be reconciled with other measurements? One possibility is that more crossbridge than ATP hydrolysis cycles occur during a transient phase after the start of unloaded shortening¹¹⁻¹³ or during steady state unloaded shortening^{8-10,27}. If there are six or more crossbridge cycles per ATP cycle, the sliding distance obtained here would be similar to that deduced from mechanical studies, 80 Å (ref. 15). The experimental data (Table 1) showed that the number of ATP hydrolysis cycles during the shortening period was about a quarter of the total number of myosin heads. Therefore, it is possible that only a fraction of crossbridges complete the ATP hydrolysis cycle during the shortening period itself. It is unlikely, however, that many of the remaining crossbridges could do work during shortening and hydrolyse ATP after shortening, because the ADP release after shortening is small (Fig. 3).

Another possibility is that when a crossbridge hydrolyses an ATP molecule and produces thin filament sliding, some free energy of hydrolysis is transferred passively to another crossbridge attached to the thin filament; that is, the first crossbridge does work on a second. Then, the second crossbridge produces sliding of the thin filament and again does work on a third. Such a cycle would occur repeatedly during one ATP hydrolysis cycle. This kind of mechanism, however, would require a very regular and directional order of crossbridge attachment-detachment cycles on the thin filament, so that the possibility may not be completely excluded but seems unlikely.

We believe a more straightforward interpretation of our results is that the range of movement over which myosin crossbridge can attach to the actin filament during one ATP hydrolysis cycle under an unloaded condition is at least 600 Å.

We thank Professor K. Maruyama and Drs A. Inoue, K. Wakabayashi and E. Prochniewicz-Nakayama for helpful discussions. We also thank Professors Th. Wieland for providing PDTMR and T. Masaki for suggesting the use of CANP.

Received 14 December 1984; accepted 14 May 1985.

- Huxley, A. F. *Prog. Biophys. molec. Biol.* **7**, 255-312 (1957).
- Huxley, H. E. *Science* **164**, 1356-1366 (1969).
- Kamiya, N. A. *Rev. Pl. Physiol.* **32**, 205-236 (1981).
- Sheetz, M. P. & Spudis, J. A. *Nature* **303**, 31-35 (1983).
- Higashi-Fujime, S. *J. Cell Biol.* **87**, 569-578 (1980).
- Huxley, H. E. *et al. J. molec. Biol.* **169**, 469-506 (1983).
- Yanagida, T. *J. Muscle Res. Cell Motil.* **6**, 43-52 (1985).
- Huxley, A. F. *Proc. R. Soc. B183*, 83-86 (1976).
- Eisenberg, E., Hill, T. L. & Chen, Y. *Biophys. J.* **29**, 195-227 (1980).
- Podolsky, R. J. & Nolan, A. C. *Cold Spring Harb. Symp. quant. Biol.* **37**, 661-668 (1972).
- Kushmerick, M. J., Larson, R. E. & Davies, R. E. *Proc. R. Soc. B174*, 293-313 (1969).
- Irving, M. & Woledge, R. C. *J. Physiol., Lond.* **321**, 411-422 (1981).
- Homsher, E., Irving, M. & Wallner, A. *J. Physiol., Lond.* **321**, 423-436 (1981).
- Wayne, A. B., Stromer, M. H., Goll, D. E. & Suzuki, A. *J. Cell Biol.* **52**, 367-381 (1972).
- Huxley, A. F. & Simmons, R. M. *Nature* **233**, 533-538 (1971).
- Wulf, E., Dehoben, A., Bautz, A., Faulstich, H. & Wieland, Th. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4498-4502 (1979).
- Yanagida, T., Nakase, M., Nishiyama, K. & Oosawa, F. *Nature* **307**, 58-60 (1984).
- Arata, T., Mukohata, Y. & Tonomura, Y. *J. Biochem., Tokyo* **82**, 801-812 (1977).
- Maruyama, K., Natori, R. & Nonomura, Y. *Nature* **262**, 58-60 (1976).
- Wang, K., McClure, J. & Tu, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3698-3702 (1979).

- Kachar, B. *Science* (in the press).
- Namba, K., Wakabayashi, K. & Mitsui, T. *J. molec. Biol.* **138**, 1-26 (1980).
- Fujime, S. *J. Physiol. Soc., Japan* **29**, 751-759 (1971).
- Wakabayashi, K. & Namba, K. *Biophys. Chem.* **14**, 111-122 (1981).
- Oosawa, F. *Polyelectrolyte* (Dekker, New York, 1971).
- Yanagida, T. & Oosawa, F. *J. molec. Biol.* **140**, 313-320 (1980).
- Podolsky, R. J. & Nolan, A. C. *Cold Spring Harb. Symp. quant. Biol.* **37**, 661 (1973).
- Tonomura, Y. *Muscle Proteins, Muscle Contraction and Cation Transport*, Chs 3, 13 (University of Tokyo Press, 1972).
- Aata, T. & Tonomura, Y. *J. Biochem., Tokyo* **80**, 1353-1358 (1976).
- Wakabayashi, K., Namba, K. & Mitsui, T. in *Contractile Mechanisms in Muscle* (eds Pollack, J. & Sugi, H.) 237-250 (Plenum, New York, 1984).
- Tregear, R. T. & Squire, J. M. *J. molec. Biol.* **77**, 279-290 (1973).
- Ishiyama, S., Sugita, H., Suzuki, K. & Imahori, K. *J. Biochem., Tokyo* **86**, 579-581 (1979).
- Ishiyama, S., Murofushi, H., Suzuki, K. & Imahori, K. *J. Biochem., Tokyo* **84**, 225-230 (1978).
- Danker, P., Low, L., Hasselbach, W. & Wieland, Th. *Biochim. biophys. Acta* **400**, 407-414 (1975).
- Yanagida, T. & Oosawa, F. *3rd Int. Congr. Cell Biol.*, 207 (1984).
- Reynard, A. M., Hass, L. F., Jacobsen, D. D. & Boyer, P. D. *J. biol. Chem.* **236**, 2277-2283 (1961).
- Hayashi, Y. & Tonomura, Y. *J. Biochem., Tokyo* **63**, 101-118 (1969).

Alterations of the *hprt* gene in human *in vivo*-derived 6-thioguanine-resistant T lymphocytes

Richard J. Albertini*‡, J. Patrick O'Neill*‡, Janice A. Nicklas*†, Nicholas H. Heintz† & Philip C. Kelleher*

Departments of *Medicine and †Pathology and ‡Vermont Regional Cancer Center, University of Vermont College of Medicine, Burlington, Vermont 05405, USA

Investigations into the extent and significance of somatic gene mutations occurring *in vivo* in humans have been hampered by the lack of a means of unambiguously defining the mutational origin of *in vivo*-derived variant cells. Several years ago we proposed that 6-thioguanine-resistant T lymphocytes, present at low frequencies in human peripheral blood, might be useful markers of *in vivo* somatic mutation^{1,2}. We and others have since described methods for the isolation and study of these unusual cells³⁻⁵. The thioguanine-resistant T cells stably lack hypoxanthine-guanine phosphoribosyltransferase (HPRT) activity, suggesting that they are somatic equivalents in normal individuals to cells from individuals with the X-chromosomal *hprt* Lesch-Nyhan germinal mutation^{6,7}. We now report that *in vivo*-derived thioguanine-resistant T-cell colonies from a single normal individual show a variety of *hprt* structural alterations, as determined by Southern blot analysis. This finding demonstrates unequivocally that these cells are genetic mutants and validates their use for fundamental and applied mutational studies in humans.

Human T cells were obtained from a healthy male donor and colonies were isolated from a clonal assay, as described by Albertini *et al.*³. Briefly, lectin-activated lymphocytes, mixed with T-cell growth factor and feeder cells, were inoculated into the wells of microtitre plates in the absence or presence of 10^{-5} M 6-thioguanine (6-TG). An average of 2 or 10^5 cells per well was inoculated into non-selected wells or 6-TG-selected wells, respectively. After 16 days, 118 of 960 non-selected wells and 24 of 960 6-TG-containing wells showed growing colonies. Cloning efficiencies, calculated by the Poisson relationship $P_0 = e^{-x}$ (where P_0 is the frequency of negative wells and x the average number of cells per well), were 6.6% and 0.025×10^{-5} in the absence and presence of thioguanine, respectively, yielding a 6-TG-resistant mutant frequency of 3.8×10^{-6} .

Twelve unselected (wild type) and 24 thioguanine-selected (mutant) colonies were isolated and subcultured; 9 unselected and 15 thioguanine-selected colonies achieved total cell numbers of $50-100 \times 10^6$ cells. All cells were positive for the surface marker T11, confirming their T-cell identities. HPRT enzyme activity⁸ in the nine unselected colonies ranged from 30 to 50 pmol inosine monophosphate produced per min per extract of 10^6 cells; <10% HPRT activity was detected in the 15 6-TG-selected colonies.

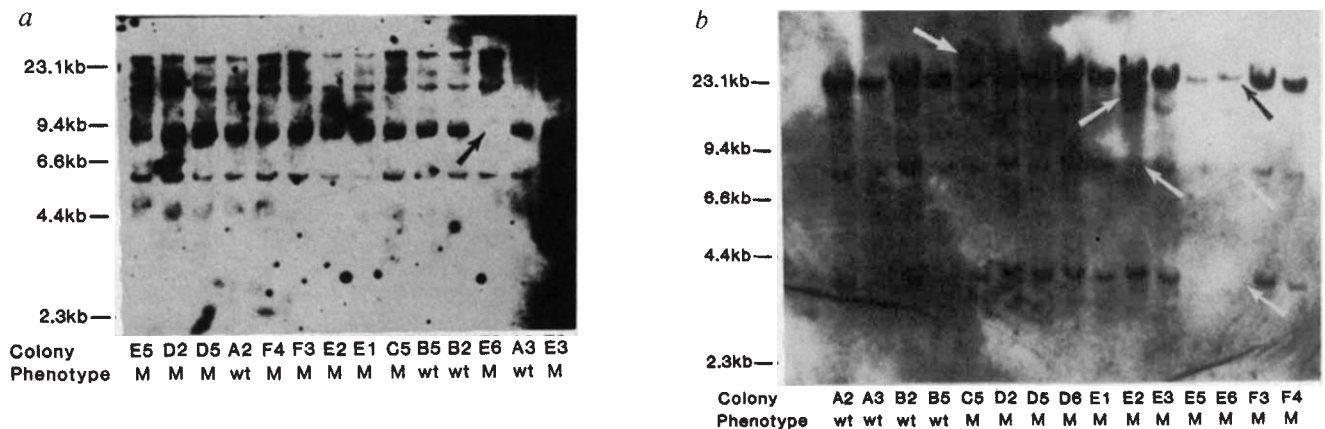


Fig. 1 Southern blots of DNA isolated from 4 (A2, A3, B2, B5) wild-type (wt) and 11 (C5, D2, D5, D6, E1, E2, E3, E5, E6, F3, F4) 6-TG-resistant mutant (M) T-cell colonies isolated from one individual and probed with a full-length *hprt* cDNA probe (pHPT31 insert⁹, given by Dr C. T. Caskey). The *hprt* insert of 947 bp contains at least 100 bp of sequence 5' to the initiation codon, the entire coding sequence and ~200 bp of the 3'-untranslated region. **a**, *EcoRI*-digested DNAs. Mutant E6 shows loss of hybridization to an 8.0- and an 8.3-kb fragment. **b**, *BamHI*-digested DNAs. Mutant C5 shows a new fragment at 35 kb, with a possible compensatory loss in the 22–25-kb region which is difficult to interpret because two *hprt* fragments and a cross-hybridizing autosomal fragment localize to this region. Mutant E2 shows a new fragment at 19 kb and loss of the 8-kb fragment. Mutant E5 shows loss of all fragments except the 25-kb autosomal fragment; however, this lane was underloaded. Mutant E6 has definitely lost hybridization to all four *hprt* fragments, although the 25-kb autosomal fragment remains. (Ethidium bromide staining and use of another, non-*hprt*-related probe demonstrated a full aliquot of DNA in all lanes except E5, clearly eliminating underloading as an explanation for fragment losses in E6).

Methods. To isolate the DNA, $15\text{--}20 \times 10^6$ frozen cells for each colony were washed and resuspended in 1 ml $T_{10}E_1$ (10 mM Tris, 1 mM EDTA pH 8.0), then 4 ml TENS (25 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM EDTA, 0.6% SDS) was added and the solution heated to 65 °C for 15 min. Proteinase K (0.5 mg) was added, and the mixture incubated at 37 °C overnight, then an additional 0.25 mg of proteinase K was added and digestion was continued for 3 h. The resulting solution was phenol-extracted twice, followed by chloroform/isoamyl alcohol (24:1) extraction, then ethanol precipitation and resuspension in $T_{10}E_1$. For the Southern blots, ~7.5 µg per lane of genomic DNA was digested with restriction enzymes, fractionated on a 0.7% agarose gel in TAE (40 mM Tris-acetate, 2 mM EDTA) buffer and transferred to nitrocellulose filters (Schleicher & Schuell). Prehybridization was for 5–16 h at 42 °C in Stark's solution¹², then $1\text{--}1.5 \times 10^6$ c.p.m. ml⁻¹ of oligonucleotide-primed ³²P-labelled *hprt* insert was added and incubation continued at 42 °C for 16 h. Washing consisted of two 10-min washes in 2 × SSC, 0.1% SDS at room temperature and two 1-h washes in 0.1 × SSC, 0.1% SDS, the first at 50 °C and the second at 60 °C. Autoradiography was performed at –80 °C with Kodak XAR-5 film and Dupont lightning plus screens for 3 days.

Southern blot analysis was used to determine alterations of the DNA in the *hprt* genes of the 6-TG-resistant colonies. Genomic DNA of 4 wild type and 11 mutant colonies was isolated from frozen samples of $15\text{--}20 \times 10^6$ cells. The DNA, cleaved with *EcoRI*, *HindIII*, *BamHI* or *MspI*, was probed with a 947-base pair (bp) full-length *hprt* complementary DNA clone⁹. This *hprt* cDNA hybridizes with four or five restriction fragments in genomic DNA, representing exons of the X-linked *hprt* gene, and also cross-hybridizes with three to four restriction fragments whose sequences have been localized on chromosomes 3, 5 and 11 (refs 10, 11).

On the *EcoRI* Southern blot (Fig. 1a), mutant E6 demonstrates a loss of probe hybridization to two fragments that migrate at 8.0 and 8.3 kilobases (kb) in the wild-type DNA. On the *BamHI* blot (Fig. 1b), mutants C5 and E2 show novel large fragments and E2 has also lost probe hybridization to a fragment at 8 kb. An *MspI* blot (Fig. 2) reveals a striking loss of fragment hybridization for mutant E6.

Using characterized subfragments of the full-length *hprt* cDNA clone, Caskey and co-workers mapped the 5'-untranslated, 5' coding, 3' coding and 3'-untranslated regions to specific restriction fragments generated by several endonucleases¹¹. Thus, it is possible to determine where alterations have occurred in the *hprt* DNA sequences of somatic 6-TG-resistant mutants by comparing the fragments absent in the mutants with the known restriction-fragment maps of the *hprt* gene^{10,11}. For the *BamHI*, *EcoRI* and *HindIII* blots, all of the missing or changed fragments reported here represent the X-linked *hprt* gene and not the autosomal cross-hybridizing sequences. However, as information on the chromosomal location and gene mapping of the *MspI* fragments has not yet been reported, definite statements on the gene alterations represented by changes in restriction fragment patterns with this endonuclease cannot be made.

Overall (see Table 1), and including data from other blots (not shown), four of the mutant colonies (D2, D6, E1 and E3) showed no restriction fragment alterations, compared with the four wild-type colonies (which, as expected, had identical patterns). Mutant E2 showed a reciprocal gain of hybridization to a novel fragment and a loss of hybridization to a wild-type *BamHI* fragment representing the 5' end of the gene. Mutant E6 demonstrated changes with all four of the enzymes tested, suggesting a deletion involving a considerable portion of the 3' end of the gene. Mutant C5 showed changes with two enzymes. *MspI* showed a gain of a large fragment, with corresponding loss of a small fragment; *BamHI* showed a gain of a 35-kb fragment with a possible loss of a 22–25-kb fragment (the migration of several fragments in the 22–25-kb region makes interpretation difficult). The blots for the remaining four mutants (D5, E5, F3 and F4) were difficult to score. Two mutants (D5 and F3) showed weak hybridization to a new 1.95-kb *MspI* fragment, perhaps as a consequence of incomplete digestion. For E5 and F4, there was insufficient DNA to test more than one or two enzymes.

The 11 6-TG-resistant mutant colonies described here were developed from progenitor mutant cells that arose *in vivo* in a single normal adult male. Thus, the different mutations do not represent a polymorphism among individuals but, rather, indicate that various molecular lesions can give rise to somatic mutations at this single locus within an individual.

Yang *et al.* demonstrates *hprt* DNA changes in 5 of 28 independent germinal Lesch-Nyhan mutations after digestion with four endonucleases¹⁰. By comparison, 3 of the 11 somatic mutants studied here (C5, E2 and E6) show obvious alterations in the *hprt* structural gene. Thus, at least four independent mutational events have occurred. However, this is certainly an underestimate for at least two reasons. First, two other colonies,

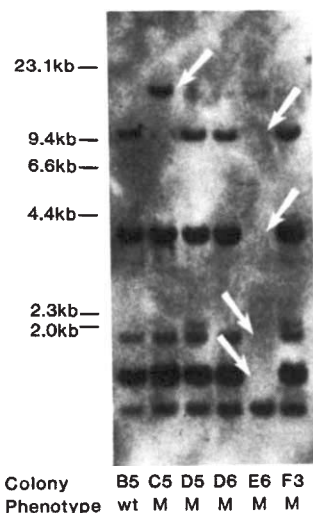


Fig. 2 Southern blot of DNA from one wild type (B5) and five thioguanine-resistant mutant (C5, D5, D6, E6, F3) colonies digested with *Msp*I and probed with the *hprt* insert from pHPT31. Mutant C5 shows hybridization to a novel fragment at 16.5 kb and reciprocal loss of the 9.4-kb fragment. Mutant E6 shows loss of hybridization to the 9.4-, 3.8-, 1.8- and 1.35-kb fragments. The new faint fragments at 2.0 kb for D5 and F3 may be a result of partial digestion and are not considered to be new fragments. (Methods as for Fig. 1).

E5 and F4, show possible structural changes with one or more of the four endonucleases used. Analysis with additional endonucleases will probably define more colonies having detectable DNA alterations. Second, we do not know what proportion of the other mutants not showing gross rearrangements were a result of small deletions, duplications or point mutations. Further study is necessary to define the number of independent mutational events responsible for this collection of thioguanine-resistant mutants. However, the results presented here do constitute unequivocal evidence for the mutational origin of the thioguanine-resistant T cells that arise *in vivo* in the normal human.

This work was supported by NIH grant R01 CA 30688 and DOE Contract US DOE AC02-83ER60147. J.A.N. is a postdoctoral trainee supported in part by training grant T32-07122-03. *Note added in proof:* A recent report¹⁴ substantiates the results reported here.

Table 1 Summary of restriction fragment changes in thioguanine-resistant mutant T-cell colonies

Mutant colonies	<i>Bam</i> HI (kb)	<i>Eco</i> RI (kb)	<i>Hind</i> III (kb)	<i>Msp</i> I (kb)
C5	+35 (–22 to 25)*			+16.5, –9.4
D2				ND
D5				
D6				
E1				ND
E2	+19, –8			ND
E3				ND
E5	(–22, –25, –8, –3.7)*	(–8.0, –8.3)*		ND
E6	–22, –25, –8, –3.7	–8.0, –8.3	–17	–1.35, –3.8 –1.8, –9.4
F3				
F4			(–40)*	

+ Indicates the presence of a novel restriction fragment; – indicates loss of hybridization to a wild-type restriction fragment. ND, not determined.

* Difficult to interpret.

Received 4 December 1984; accepted 8 May 1985.

1. Strauss, G. H. & Albertini, R. J. *Mutat. Res.* **61**, 353–379 (1979).
2. Albertini, R. J. in *Indicators of Genotoxic Exposure, Banbury Rep. 13* (eds Bridges, B. A., Butterworth, B. E. & Weinstein, I. B.) 393–410 (Cold Spring Harbor Laboratory, New York 1982).
3. Albertini, R. J., Castle, K. & Borcharding, W. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6617–6621 (1982).
4. Strauss, G. H. in *Indicators of Genotoxic Exposure, Banbury Rep. 13* (eds Bridges, B. A., Butterworth, B. E. & Weinstein, I. B.) 423–441 (Cold Spring Harbor Laboratory, New York, 1982).
5. Morley, A. A., Trainor, K. J., Seshadri, R. & Ryall, R. B. *Nature* **302**, 155–156 (1983).
6. DeMars, R. *Fedn Proc.* **30**, 944–945 (1971).
7. Seegmiller, J. E., Rosenbloom, F. M. & Kelley, W. N. *Science* **155**, 1682–1684 (1967).
8. Fuscoe, J. C., O'Neill, J. P., Machanoff, R. & Hsie, A. W. *Mutat. Res.* **96**, 15–30 (1982).
9. Brennand, J., Konecki, D. S. & Caskey, C. T. *J. biol. Chem.* **258**, 9593–9596 (1983).
10. Yang, T. P. *et al. Nature* **310**, 412–414 (1984).
11. Patel, P. I. *et al. Somat. Cell molec. Genet.* **10**, 483–493 (1984).
12. Wahl, G. N., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
13. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **137**, 266–267 (1984).
14. Turner, D. R., Morley, A. A., Haliandros, M., Kutlaca, R. & Sanderson, B. J. *Nature* **315**, 343–345 (1985).

A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity

Valakunja Nagaraja, John C. W. Shepherd & Thomas A. Bickle*

Microbiology Department, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Early attempts to generate new restriction specificities by recombination between allelic restriction-modification systems have been unsuccessful¹. Bullas *et al.*² succeeded in isolating a new specificity, SQ, in *Salmonella* that they interpreted as being the result of a recombination event between the parental strains, *Salmonella typhimurium* and *S. potsdam*, which encode the SB and SP restriction systems, respectively. This interpretation has recently been confirmed by DNA heteroduplex studies with the SB, SP and SQ structural genes³. We have determined the DNA sequences recognized by the SB and SP enzymes⁴ and found that, like all type I restriction sequences, they are split into two specific domains by a spacer of nonspecific sequence that, for both SB and SP, is 6 base pairs (bp) long. We have now determined the sequence recognized by the recombinant SQ enzyme and find that it is a hybrid between the SB and SP sequences, containing one specific domain from each parental strain. This result implies that each of the two specific domains is recognized by a physically distinct part of the enzyme.

The SQ enzyme was purified from the strain L4004 (ref. 2) by methods used previously for SB and SP⁴. Its recognition sequence was determined using a computer-aided strategy described in detail elsewhere⁴. The method depends on the fact that type I restriction enzymes are also modification methylases and thus can be used to introduce radioactive methyl groups into their recognition sites in sequenced DNA molecules. The radioactive label is then mapped to small regions of the DNA by restriction analysis and fluorography and computer analysis is used to find DNA sequences common to all labelled regions and absent from all others.

We screened various different sequenced DNA molecules for the presence of SQ sites by *in vitro* methylation with [³H-Me]-S-adenosyl methionine. DNA molecules that had no SQ sites were pBR322, pBR325, pSH144, pFR100, ØX174, G4, M13 and simian virus 40. Phage λ DNA, which has been reported not to be restricted by the SQ system², was a substrate for the enzyme. In our hands SQ does restrict, with plating efficiencies ranging from 10^{–1} to 10^{–2} in different experiments. Mapping experiments showed that λ DNA contains a single site for the enzyme, which

* To whom correspondence should be addressed.

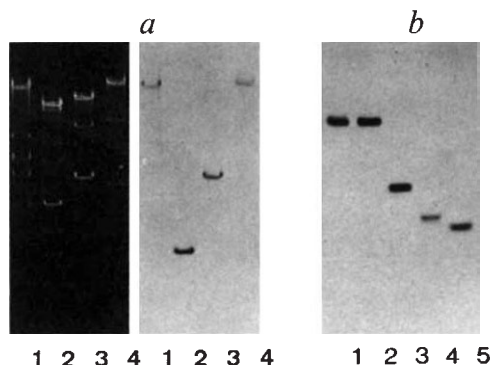


Fig. 1 Mapping of the SQ site in phage λ DNA. The DNA was methylated with purified SQ enzyme and [$^3\text{H-Me}$]-S-adenosyl methionine (Amersham, 62-74 Ci mmol $^{-1}$) as described earlier⁴. The labelled DNA was then cleaved with different type II restriction enzymes and the fragments were separated on either a 1% agarose gel (a) or an 8% polyacrylamide gel (b). The gels were photographed after staining with ethidium bromide and then fluorographed. a, Mapping of the SQ site with enzymes that cut the DNA infrequently. The left panel shows the stained gel, the right panel the fluorogram. The enzymes used and the extent of the labelled fragments are: lane 1, *EcoRI*, 1-21,226; lane 2, *AccI*, 1-2,190; lane 3, *AvaI*, 1-4,720; lane 4, *HindIII*, 1-23,120. b, Fine mapping of the site and confirmation of the sequence. Only the fluorogram is shown. Lane 1, *HaeIII*, 403-469; lane 2, *HaeIII* + *RsaI*, the same fragment is labelled as in lane 1, even though *RsaI* has a site at position 439; lane 3, *BstNI* + *HaeIII*, 425-469; lane 4, *BglII* + *TaqI*, 415-451; lane 5, *BstNI* + *FnuDII*, 425-459. We conclude that a SQ site is present between positions 425 and 451.

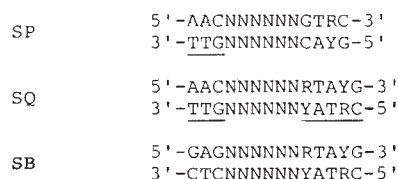


Fig. 2 The recognition sequences of SP, SB and SQ. Homologies between the sequences are indicated by bars.

we first mapped to the left 2,190 bp of the DNA molecule (Fig. 1a) and then, using combinations of fine-cutting enzymes, to a 26-bp stretch extending from positions 425 to 451 of the λ DNA sequence⁵ (Fig. 1b). Another SQ site was found and mapped to a small region of the yeast *PHO5* gene⁶ (not shown). The two regions that contained SQ sites, together with all the sequences that did not, were subjected to computer analysis as described above. Most sequences were taken from the DNA sequence databank of the European Molecular Biology Laboratory (EMBL). This search failed to find a non-degenerate sequence that could be a candidate for the SQ recognition site, but further searches found a single sequence with one degeneracy that fitted the data 5'-AACNNNNNNRTACG-3', where N is any nucleotide and R is either purine. The experiment shown in Fig. 1b, lanes 1 and 2, confirms this sequence. In phage λ , G is present in the position indicated as R and the resulting 5'-GTAC-3' is an *RsaI* recognition site. Figure 1 shows that *RsaI* does not cut this site in SQ-modified DNA, indicating that SQ has methylated the DNA within the site.

We then searched the EMBL DNA sequence databank for other DNA molecules containing the putative SQ site and variants of it and found that the plasmids p3201 (ref. 7) and colE1 (ref. 8), phage T7 DNA⁹ and the yeast *CYTcI* (ref. 10) gene were suitable for further analysis. The SQ sites in all of these DNA molecules were labelled and mapped (not shown) and the results are summarized in Table 1. One further degeneracy in the sequence was found and the SQ recognition site is 5'-AACNNNNNNRTAYG-3', where Y can be either pyrimidine. Sequences that had pyrimidines in the position marked by R or purines in the position marked by Y were not

Table 1 Compilation of SQ recognition sequences

DNA	Position	Sequence
λ	430	GAAGTATTGAGTACGA
PHO5	1,612	TAAGTCCGAATACGCT
CYTcI	1,401	AAACAATTGAGTATGC
ColE1	3,439	CAACAGGACTATATGC
T7	19,953	GAAGTGGACCATATGT
T7	31,459	CAACGTTGATGTACGC
p3201	1,142	TAAGTGGGTATATGT
p3201	2,495	GAAGTTGATGGTATGC
Consensus:		AACNNNNNNRTAYG

substrates for the enzyme. All four nucleotides can be found flanking the sequence and at all positions of the nonspecific spacer (Table 1).

Figure 2 compares the recognition sequence of SQ with those of its parents, SB and SP. SQ has the AAC domain of SP and the RTAYG pentanucleotide domain of SB. It is clear that the recombination event that generated SQ took place within the region encoding the part of the polypeptide that is involved in sequence recognition.

The allelic type I restriction enzymes found in *Escherichia coli* and *Salmonella* spp. are complex enzymes, and DNA recognition is a function of just one of their subunits, the product of the *hdsS* gene (reviewed in ref. 11). Several *hdsS* genes have been sequenced¹² and the sequences are surprisingly non-homologous, only two short regions, one in the middle of the gene and the other at the 3' end, being highly conserved. The recombination that yielded SQ took place in the central conserved region³. It is not known whether the parts of the polypeptide that bind the recognition sequence are encoded by this region or by the nonconserved sequences on either side.

Recombination between *hdsS* genes has the potential for generating any combination of specific sequence domains and thus of creating new sequence specificities. There is some indication that this may have happened naturally in the past. For example, *EcoK*, SP and SQ all share a common AAC domain while GAG is common to both *EcoB* and SB. One response of bacteriophages to the presence of restriction enzymes in their hosts is to mutate the recognition sites in the phage genomes, making them resistant to restriction^{13,14}. It is therefore advantageous for the host to be able to change the recognition specificity of its restriction system from time to time.

We thank P. G. Baas, A. Hinnen, H. S. Jansz, C. Sengstag, D. Sherratt, F. W. Studier and A. P. G. M. Van Loom for gifts of phages and plasmids, C. Price for helpful comments on the manuscript and T. Pripl for help with the experiments. This work would not have been possible without access to the EMBL DNA sequence databank, and was supported by grants from the Swiss NSF.

Received 14 March; accepted 31 May 1985.

- Arber, W. & Linn, S. A. *Rev. Biochem.* **38**, 467-500 (1969).
- Bullas, L. R., Colson, C. & Van Pel, A. J. *gen. Microbiol.* **95**, 166-172 (1976).
- Fuller-Pace, F. V., Bullas, L. R., Delius, H. & Murray, N. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6095-6099 (1984).
- Nagaraja, V., Shepherd, J. C. W., Pripl, T. & Bickle, T. A. *J. molec. Biol.* **182**, 579-587 (1985).
- Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. *J. molec. Biol.* **162**, 729-773 (1982).
- Bajwa, W., Meyhack, B., Rudolph, H., Schweingruber, A. M. & Hinnen, A. *Nucleic Acids Res.* **12**, 7721-7739 (1984).
- Sengstag, C. & Arber, W. *EMBO J.* **2**, 67-71 (1983).
- Chan, P. T., Ohmori, H., Tomizawa, J.-I. & Lebowitz, J. *J. biol. Chem.* (in the press).
- Dunn, J. J. & Studier, F. W. *J. molec. Biol.* **166**, 477-535 (1983).
- Sadler, I., Suda, K., Schatz, G., Kaudewitz, F. & Haid, A. *EMBO J.* **3**, 2137-2143 (1984).
- Bickle, T. A. in *The Nucleases* (eds Linn, S. & Roberts, R. J.) 85-108 (Cold Spring Harbor Laboratory, New York, 1982).
- Gough, J. A. & Murray, N. E. *J. molec. Biol.* **166**, 1-19 (1983).
- Arber, W. & Kuhnlein, U. *Path. Microbiol.* **30**, 946-952 (1967).
- Kruger, D. H. & Bickle, T. A. *Microbiol. Rev.* **47**, 345-360 (1983).

The laser — hero or villain?

from Bridget Marx

Although too often associated in people's minds with star wars weaponry the laser has become an invaluable tool in many varied and peaceful applications.

In the early days the laser, for Light Amplification by Stimulated Emission of Radiation, was dubbed 'a solution looking for a problem', but in the last quarter of a century lasers have found application in areas as diverse as materials processing and the treatment of cancer.

The medical application of lasers is now in fact a major industry—out of an estimated total commercial laser market of about \$400 million, at least one quarter is connected with medical and surgical applications. This market is growing at a rate of more than 30% per year and has become such a significant factor for the manufacturers of lasers that several companies have actually been bought out by medical conglomerates.

Laser surgery

The laser was first demonstrated in 1960 by Theodore Maiman at the Hughes Research Laboratories in California, and the first 'medical' application came when Maiman's ruby laser was being used in animals' eyes, as a source for photocoagulation and by 1962 human patients were being treated. As new lasers became available these were quickly tried out in the medical field. The argon ion laser was first considered in 1965, the clinical trials were started in 1969 and the first commercial systems became available in 1971. The argon laser has proved a lot more successful for the treatment of vascular diseases of the retina than the original ruby systems and there are now many thousands of medical argon lasers worldwide. At the time of writing lasers in routine use in hospitals include the argon ion laser, the carbon dioxide laser, the Nd:YAG laser and, to a lesser extent, the krypton ion laser. Those at present in the early stages of development include the excimer,

gold vapour and copper vapour lasers.

The laser is in many ways ideally suited to medical applications, providing as it does a light source that is powerful enough to destroy the target tissue yet is controllable in terms of power, spot size and wavelength. The light can be easily and efficiently taken to the required site either directly or by optical fibre. These advantages led to the rapid development of laser surgery in dermatology, ophthalmology and gynaecology. Other exciting areas are also opening up, such as photoradiation therapy for the treatment of cancer, laser recanalization of clogged arteries, laser welding of soft tissues, and even laser acupuncture for pain relief.

Diagnostics

The direct surgical use of lasers is not the only clinical application of lasers, and after photocoagulation, the 'surgical' application which accounts for 40% of all water-cooled lasers in use, the makers of cytofluorescence equipment are among the biggest consumers of water-cooled ion lasers, accounting for a further 14% of the market. Lasers allow biological specimens to be analysed for their cell content and the different cells to be sorted into different containers depending on their light scattering characteristics. This method of detecting cancer cells is very straightforward and cell-sorting equipment is to be found in many hospital laboratories.

Non-medical

The largest market for lasers is still in academic and industrial research and development: nearly \$100 million out of the total \$400 million. A traditional area of pure research is spectroscopy, and other growth areas are laser Doppler velocimetry and holography. Holography

these days is used not just for making pretty pictures, but also for making non-forgeable security passes and credit cards. And even spectroscopy has left the realms of pure research as we move into the age of laser isotope separation. (Laser isotope separation accounts for more than one eighth of the total CW dye laser market each year.) There is also a large part of the world's laser research effort concentrated in the field of optical-fibre communications, so that in the future telephone messages will not be relayed by an electrical signal along a wire but will be carried by a modulation of a laser beam that is being conducted along an optical fibre.

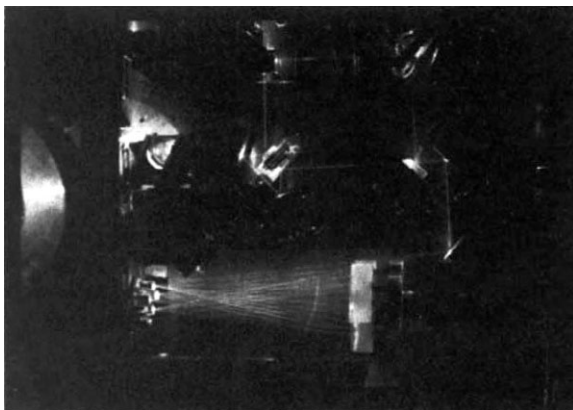
Popularization

Apart from these scientific applications there is a plethora of relatively 'low-tech' areas where lasers are carving out a niche. The entertainment industry is a market sector worth several millions of dollars each year. Laser light shows are quite commonplace these days in discotheques or at big public gatherings, as demonstrated during the 1984 Olympic Games in Los Angeles. There are also schemes using lasers to project television images onto a large screen with enough brilliance to be seen in daylight. And there are markets that use large volumes of low-cost lasers. In the entertainment industry the home video and audio disk systems come into this category. There are also laser printers and bar-code reading scanners at supermarket checkouts: these account for about 10% of the laser market, using approximately 100,000 small lasers a year.

In heavier industries too, lasers are gaining ground. Highpower infrared lasers are used for cutting, drilling and welding in laser machine tools. Ultraviolet excimer lasers are used in photolithography and other developmental techniques in the semiconductor industry.

So how will history judge the laser as it emerges from its formative years: Hero or villain? It has proved its worth in the areas of scientific research and manufacturing. It is capable of many medical feats, from miracle cure to routine surgery. President Reagan's Strategic Defense Initiative is based on the hope that it has an equally destructive side to its nature. □

The largest market area for lasers is still research and development. One application in which basic physics is tested is shown here. In this exotic application an argon ion laser is used to illuminate a long baseline interferometer for the detection of gravitational radiation. Interferometers with baselines of many kilometres are planned. The photograph shows an early prototype Michelson interferometer with a relatively short baseline. (Photo courtesy of Glasgow University.)



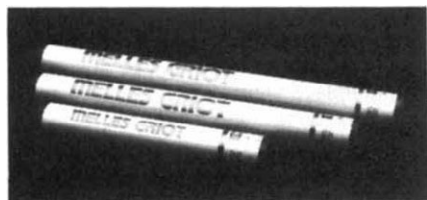
Bridget Marx is UK Sales Manager of the Laser Products Division of Coherent (UK) Ltd, Cambridge Science Park, Milton Road, Cambridge CB4 4BH, UK.

Lasers in commercial applications

A user-group based at Culham gives access to high-power lasers, and laser Raman microscopy is becoming established.

• A series of five reports from the Laser Analytics Division of Spectra-Physics should be of interest to all those using mass spectroscopy for the determination of trace moisture in hermetically-sealed integrated circuit packages. The reports show that the results of the infrared testing of IC packages sealed under controlled moisture conditions exhibit significant gaseous and absorbed moisture discrepancies when compared to less-accurate mass spectroscopy. Titles of the five are: (1) 'Water Vapor Measurements in IC Packages Using an IR Diode Laser'; (2) 'Trace Analysis of Moisture in IC Packages Using Derivative Diode Laser Spectroscopy'; (3) 'Laser Spectrometric Determination of Moisture in IC Packages'; (4) 'Dynamic Measurement of the Water Vapor Content of IC Packages Using Derivative IR Diode Laser Spectroscopy'; (5) 'Trace Analysis of Microvolume Gas Samples Using IR Spectroscopy Moisture in IC Packages'.

Reader Service No. 100.



The new Melles Griot laser is an alternative to solid-state systems.

• A new helium/neon laser with infrared output at $1.523 \mu\text{m}$ is now available from Melles Griot. The new laser matches the output of solid-state laser diodes, without the inherent problems, and should find applications in data transmission, intrusion alarms, target designation, range finding and security systems. Current models offer 0.25, 0.5 or 1.00 milliwatts of TEM₀₀ single longitudinal mode collimated coherent output. This output matches the minimum attenuation in the new high performance OVD fibres, falls in the middle of a clear atmospheric window, corresponds to the sensitivity peak of InGaAs detectors and is classified 'eye safe' by the US Bureau of Radiological Health. These $1.5 \mu\text{m}$ He/Ne lasers are constructed with zero-leak hardseals for longevity, and feature a patented symmetrical cold-cathode for stability. The instruments are based on the same design and manufacturing techniques as the company's established 632.8nm range.

Reader Service No. 101.

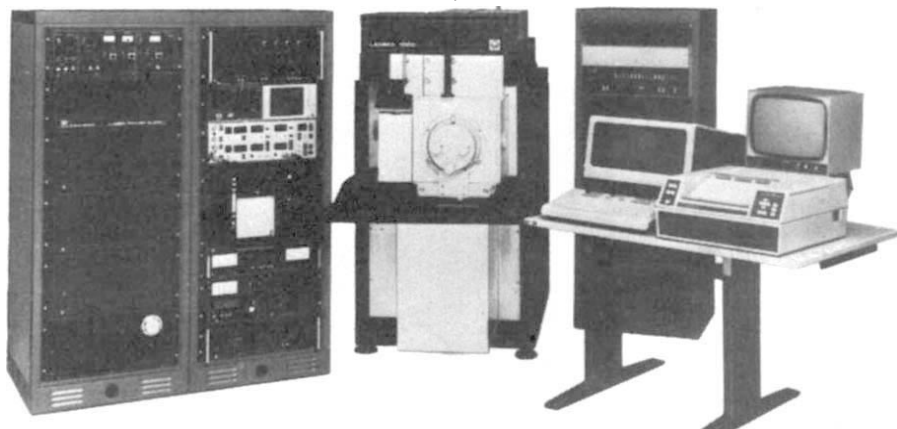
These notes are based on information provided by the manufacturers. To obtain further details about these products use the reader service card bound inside the journal.

• The Hybond Laser Spotlight Bonder, specially developed for wire-bonding, features enhanced accuracy provided by a laser spotlight which enables the work piece to be positioned quickly and with great precision. An integral 'soft-touch' ramping force generator protects the work piece against thermal shock. Precise control of position is achieved by a combination of easy-to-use adjustable controls for search, bond and loop height, with a motorized z function for true vertical movement. Quick visual reference and ease of fine adjustment is provided by the 'Paramagraph' controls. In the United Kingdom the bonder is available from Intertrade Scientific of Milton Keynes.

Reader Service No. 102.

• The Leybold Heraeus LAMMA 1000 uses a Nd:YAG laser to provide approximately 2mJ of light energy in 15 ns focused to $1 \mu\text{m}$ spot size. This laser power is used to ionize a micro-volume of the sample which is held on a precision x-y-z table in the sample chamber. The ions formed are then accelerated into a time of flight drift tube and are mass resolved. The ion detector at the end of the drift tube has a copper-beryllium dynode electron multiplier giving an amplification of better than 10^6 . The ease of putting a sample into the instrument with the minimum of sample preparation makes this instrument particularly useful for analysing small bulk samples. The lateral resolution allows the determination of distribution of atomic or molecular species. It is also possible to trace stable isotopes and to do depth profiling with LAMMA. By reducing the laser power, larger ionic fragments are produced for analysis of organic substances. It is possible to obtain positive or negative spectra and masses greater than 1,000 amu can be resolved. LAMMA can detect beryllium, lithium, helium and hydrogen.

Reader Service No. 103.



The LAMMA system from Leybold Heraeus resolves masses greater than 1,000 amu and easily detects beryllium, lithium and hydrogen.

• Culham Laboratory's Laser Application Group produces two multi-kilowatt lasers, the CL5 and CL10, together with the associated beam-manipulating and work-handling facilities. Applications include material processing, such as cutting, welding and surface treatment, high power CO₂ laser beam propagation, monitoring and quality control. The Laser Applications Group, which is situated at the UK Atomic Energy Authority's Culham Laboratory, carries out contract research and development on high power laser systems for customers from both industry and government-funded organizations. The group has assembled an extensive range of CO₂ lasers, with output powers from a few watts up to greater than 10kW. A wide variety of beam-manipulating and work-piece handling equipment has also been developed and experience has been gained in the propagation of multi-kilowatt laser beams over distances greater than 50m. Reliable techniques have been developed for materials processing and for analysing the laser beam quality. This expertise enables the Laser Applications Group to offer customers services from a short assessment to the provision of a complete laser-based flexible manufacturing system. Development programmes are undertaken either for individual customers or on a group-sponsored basis. A 'Laser Users' Club' has been formed at Culham to provide the member companies with the necessary knowledge of laser processing technology for them to remain competitive. The Laser Club programme, has been sponsored jointly by eight companies and the Department of Trade and Industry. The Laser Applications Group aims to establish other club activities, particularly with European participation, and these could be directed towards specific industries.

Reader Service No. 104.



Rofin-Sinar's optical spectrum analyser for the range 200–4,700 nm.

• Rofin-Sinar Laser UK Ltd has broadened its range of optical spectrum analysers with the introduction of the Model RSO 6230 spectral processor. The new unit is fully computerized and includes an IEEE and RS232 interface. The system includes a 9-inch computer screen, a keyboard, two 160K, 3½-inch industrial-type floppy disk drives, integrated computer and all controls needed to interface directly to the RSO 6100 series monochromator heads. It can provide the user with a wavelength and amplitude corrected spectrum. A full scan, sampling the spectrum every 0.5nm, takes 1.6 seconds. Alternatively if the spectrum is sampled every 10nm, a complete spectrum can be acquired in 100ms. There are interfaces for a printer, IEEE, RS232 and oscilloscope built into the unit, and a semiautomatic calibration feature provides speed and ease of wavelength calibration. Wavelength accuracy is typically better than 1nm. The spectral range may be covered from 200 to 4,700nm.

Reader Service No. 105.

• High performance photon drag detectors, used for the temporal analysis of high peak power CO₂ laser pulses, are now manufactured by Edinburgh Instruments of Riccarton, Scotland. These detectors are available in two formats for use as either an in-line monitor with 75% transmission, or an end of the line detector. Both types are characterized by the usual well-established properties of photon drags — room temperature operation, very fast response (less than 1ns), high damage threshold, compactness and ease of use. They also offer extra wide-angle access for easy alignment, a stainless steel housing, elimination of photodiode effects at the electrode, an insulated mounting sleeve to significantly reduce the transient noise pick up from TEA lasers and good long-term stability. These detectors have a guaranteed spatial uniformity of $\pm 2\%$ compared with conventional photon drag detectors which typically exhibit a sensitivity at the edge of the face up to 40% higher than at the centre. Both of Edinburgh Instruments' photon drags have a usable aperture of 10mm and the typical responsivity of the monitor is 180 mV per MW and the detector is typically 200 mV per MW. Damage threshold figures are 20 MW per cm².

Reader Service No. 106.

• Digital delay generators are widely used in pulsed laser timing, fast circuit analysis and fibre optic testing to accurately and repeatably control time. A signal is used to trigger the generator which produces a pair of output pulses, an initial pulse and a delayed pulse. The duration between the two pulses is digitally selectable according to the requirements of the experimenter. The new Model 7695 from Berkeley Nucleonics Corporation has been designed to make gate and delay generation in the sub-nanosecond region not only possible but also highly repeatable. Delays of up to 10ms can be dialled into the unit in 50 ps steps. Until now delays of such short duration have not been possible without problems connected with accuracy, jitter and transition times. The BNC Model 7095 has a guaranteed jitter of less than 25 ps r.m.s. maximum, and an accuracy of ± 0.1 ns on delay generation, up to 9,950 ps. Various options are available including full IEEE 488 compatibility which includes an internal RAM for storing up to sixteen settings. The extremely short delay times attainable with the Model 7095 make it suitable for time of flight experiments, radar simulation, shock-wave physics, chemical kinetics, plasma physics, fibre optics and laser applications.

Reader Service No. 107.



Beam Scan for the analysis of laser output.

• Oriel Scientific's Beam Scan system is a compact analyser for the comprehensive characterization of the output parameters of lasers and imaged white light sources. Available in several models spanning the 0.1 to 1,000 μ m range, Beam Scan can measure beam size, spatial profile, ellipticity, divergence and collimation. The instrument, which utilizes a moving narrow slit (or pinhole) and a photosensitive detector, continuously samples the energy of a beam incident on it, measuring the intensity profile and relaying, for example, the beam size, at a pre-selected intensity level (for example $1/e^2$) to the front panel LED display. Alternatively this information is available, from an output jack, as a direct analogue signal for display on an oscilloscope. For optical alignment, focal length and image quality checking, both in the visible, and in the non-visually observable infrared, Beam Scan can quantify or track beams from 7.5 μ m to 17mm in diameter. Among its application areas, Beam Scan can be used to characterize fibre optic sources, mode filling and numerical aperture, or compare acousto-optic input and output beam parameters.

Reader Service No. 108.

ADVERTISEMENTS

NEWAT '85

NEWDAT data base now available on 5¼" floppy disks for IBM PC's and look-alikes. More than 1200 amino acid sequences with literature citations. Six-disk kit includes programs for entering data, searching and aligning. \$250. Limited supply.

Telephone: 619-453-2467

or
NEWAT
P.O. Box 12822,
La Jolla
CA 92037 USA

Reader Service No.15

LIPOSOMES

REPRODUCIBLE – HOMOGENEOUS –
EXTREMELY RAPID

Our new commercial LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized

- UNILAMELLAR OR
- OLIGOLAMELLAR OR EVEN
- MULTILAMELLAR LIPOSOMES

depending on the used experimental conditions. Liposome size is selectable between 25 and ca. 600 nm diameter up to 1000 nm for multilamellar liposomes. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. High encapsulation efficiency for hydrophilic and/or lipophilic agents. Lipid concentration can be up to 300 mg/ml. Lipid loss during preparation is extremely low: down to 0.1%.

DIANORM - Geräte

POB 126, D-8000 Munich 65, FRG

Reader Service No.13

NOVEL 1, 2-DIACYLGLYCEROLS AND PHORBOL ESTERS FOR PROTEIN KINASE C ACTIVATION AND CANCER RESEARCH

Nine 1,2-DGs, including OAG, MAG DiC:8 and 1-Stearoyl-2-arachidonyl-glycerol for PKC activation, PMA(TPA) and PDB for Tumor promotion and several other high purity Lipids, Nucleotides and Bio-chemical reagents available from stock at unbeatable prices and quality.

Call (414) 355-9580 for details
or Write to

Life Science Resources
P.O. Box 23021, Milwaukee
WI 53223

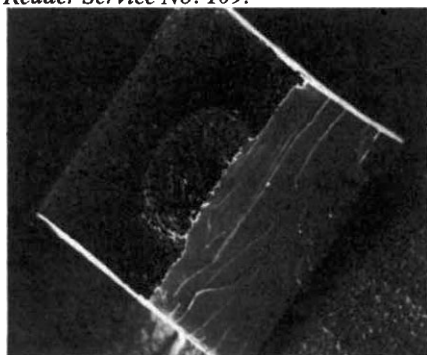
Reader Service No.14

nature
is available
in microform
from University
Microfilms
International.

Call toll-free 800-521-3044. In Michigan, Alaska and Hawaii call collect 313-761-4700. Or mail inquiry to: University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106.

• New from Labsphere is an infrared integrating sphere for CO₂ laser power measurements which uses a 4-inch diameter water-cooled sphere with diffuse gold coating. The WIS-4000 can be used in conjunction with almost any infrared detector which mounts at a port on the integrating sphere. The high power levels of infrared lasers are attenuated to a 'safe' value within the operating range of the detector and alignment sensitive measurement errors are eliminated. The WIS-4000 can also be mounted at the back mirror of the laser cavity to monitor actual power levels during operation. The WIS-4000 is normally supplied with a gold coating optimized for diffuse reflectance at the CO₂ laser wavelength. However, the sphere may be supplied optimized at other infrared laser lines as well. Accessories for the WIS-4000 include a water-cooled laser target used to reflect the first strike of the laser and a constant temperature recirculating cooler.

Reader Service No. 109.



New from Spectra Physics, a single-mode tunable diode laser.

• The new range of short-cavity lead-suit lasers from the Laser Analytics Division of Spectra Physics provides substantially improved single-mode performance at liquid nitrogen temperature and above. The latest lasers make possible the achievement of single-mode, CW emission from short-cavity mesa-stripe geometry Pb-salt diode lasers in 5, 10 and 30 μm spectral regions. The new devices, with a cavity length of 120 μm exhibit a Fabry-Perot mode separation of 6 cm^{-1} . Exploiting the spectral narrowing of the optical gain characteristics with increasing operating temperature single-mode operation can be obtained over broad current ranges in the 5 and 10 μm spectral regions at liquid nitrogen temperature and above. For the 30 μm lasers, single-mode operation is possible in short-cavity lasers at current values of up to almost twice threshold. Single-mode tunable diode laser (TDL) operation in the mid-infrared spectral region is of great interest in applications such as high-resolution infrared spectroscopy, heterodyne detection, and, in the future, mid-infrared fibre transmission. In addition, operating such lasers at liquid nitrogen temperature and above greatly simplifies the cooling requirements.

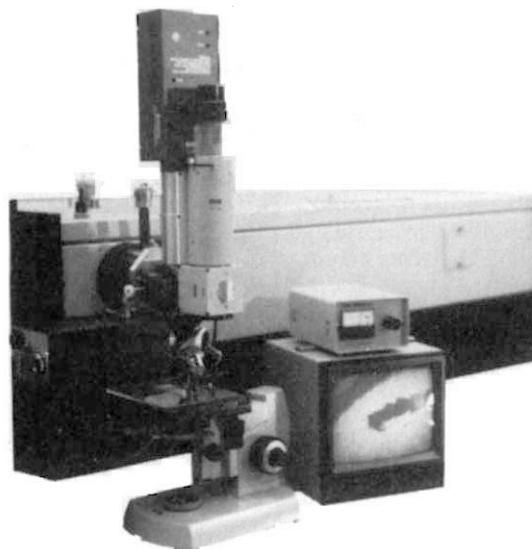
Reader Service No. 110.

• A series of Laser Beam Analysers (LBA) manufactured in West Germany by ALL is now available in the United Kingdom. The LBA accurately monitors the spatial distribution of a laser beam intensity in two orthogonal axes. The instruments use a highly reflective molybdenum needle rotating at 50Hz perpendicular to the beam path that reflects a fraction of the beam onto two pyroelectric detectors. The amplitude of the signals from these detectors displayed simultaneously on a dual trace oscilloscope corresponds to the power distribution through the beam. The LBA can handle beam powers from 10W to 20KW, with power densities up to 10 W cm^{-2} . As the power loss of the beam is only 0.8% as it passes through the LBA, the instruments can be used for on-line monitoring of beam profiles during experiments and material processing. The LBA can monitor changes in the mode before they become apparent in the quality of the laser processing. Other applications include measurement of beam divergence and beam focus by movement along the laser beam path.

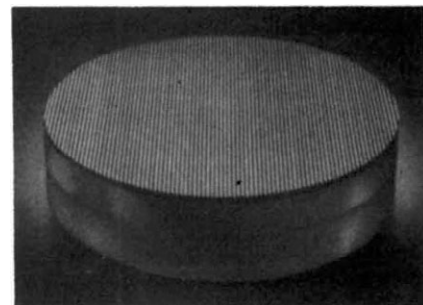
Reader Service No. 111.

• Laser Raman microscopy is now established as a powerful analytical technique. It provides a means of examining and chemically analysing samples as small as 2 μm in size in a non destructive way. The laser beam, often with a power of less than 100 mW, enters the microscope and is focused on the area of the sample under investigation. The Raman scattered light is collected and enters the spectrometer which then produces the vibrational spectrum of the sample. The technique has proved invaluable for studying optical fibres, semiconductor films, lubricants, corrosion effects, fluid inclusions in geological samples, biological tissues, waste products and polymers. Application notes are available on most of these applications from Glen Creston Instruments.

Reader Service No. 112.



Glen Creston Instruments' laser Raman microscope can analyse samples only 2 μm in size in a non-destructive manner. Applications include study of fluid inclusions in geological samples as well as analysis of biological tissues.



The Littrow-mounted diffraction grating end reflector.

• The Optometrics Group, long established in the field of photonics, has introduced the model LA-1000, the latest in a series of light harnessing instruments. This infrared Laser Attenuator gives researchers greater flexibility in experiments requiring precision attenuation of infrared laser beams. The LA-1000 is primarily designed for linearly polarized lasers, especially CO₂ lasers at 10.6 μm . For unpolarized lasers, it is possible to use two attenuators in series. Also from Optometrics is a new system for wavelength tuning and selection. The output wavelength of a laser can be tuned by turning the Littrow-mounted diffraction grating end reflector about an axis parallel to the grooves. The equation relating wavelength to angle of incidence on the grating is $n\lambda = 2d \sin i$. Using this relationship, one can calibrate a mirror mount to i using a grating as one of the reflectors in the laser cavity. This will narrow the spectral line width of the output in a region around the Littrow wavelength.

Reader Service No. 113.

• Spectra Diode Labs has introduced a new range of 200-mW continuous wave diodes. Claimed to be the most powerful diode lasers commercially available, they should find a niche in such fields as solid-state lasers, and fibre optic power delivery.

Reader Service No. 114.

Mathematics

Applied Exterior Calculus. By DOMINIC G.B. EDELEN. Wiley: 1985. Pp.471. ISBN 0-471-80773-7. £56.95.

Cluster Dissection and Analysis: Theory, Fortran Programs, Examples. By HELMUTH SPÄTH. Horwood: 1985. Pp.226. ISBN 0-85312-736-0. £25.

Differential Algebraic Groups. E.R. KOLCHIN. Academic: 1985. Pp.217. ISBN 0-12-417640-2. \$60, £46.

Elements of Algebra. By L. EULER. Springer-Verlag: 1984. Pp.593. ISBN 3-540-96014-7/0-387-96014-7. DM88, \$30.90.

Graphs and Applications: Proceedings of the First Colorado Symposium on Graph Theory. By FRANK HARARY and JOHN S. MAYBEE. Wiley: 1985. Pp.347. ISBN 0-471-88772-2. £56.95.

An Introduction to Probability Theory with Statistical Applications. By MICHAEL A. GOLBERG. Plenum: 1984. Pp.662. ISBN 0-306-41645-X. \$69.50.

Learning the Art of Mathematical Modelling. By M. CROSS and A.O. MOSCARDINI. Ellis Horwood: 1985. Pp.155. ISBN 0-85312-780-8. £7.50.

The Mathematical Basis of Finite Element Methods. DAVID F. GRIFFITHS (ed.). Oxford University Press: 1985. Pp.189. ISBN 0-19-853605-4. £20.

Minimization Methods for Non-Differentiable Functions. By N.Z. SHOR. Springer: 1985. Pp.162. ISBN 3-540-12763-1. DM84.

Numerical Solution of Elliptic Problems. By GARRETT BIRKOFF and ROBERT E. LYNCH. SIAM/Wiley: 1984. Pp.319. ISBN 0-89871-197-5. £30.

Progress in Graph Theory. J. ADRIAN BONDY and U.S.R. MURTY (eds). Academic: 1984. Pp.539. ISBN 0-12-114320-1. \$59.50, £46.

Seminar on Nonlinear Partial Differential Equations. S.S. CHERN (ed.). Springer-Verlag: 1984. Pp.373. ISBN 3-540-96079-1/0-387-96079-1. DM94.

Cosmology

Building the Universe. CHRISTINE SUTTON (ed.). Basil Blackwell: 1985. Pp.361. Pbk ISBN 0-631-14103-0. Hbk £19.50; pbk £7.95.

High Energy Astrophysics and Cosmology. J. YAND and C. ZHU (eds). Gordon and Breach: 1983. Pp.527. ISBN 0-677-3134-03. \$65.

Physics of Thermal Gaseous Nebulae, Vol. 112. By LAWRENCE H. ALLER. Reidel: 1984. Pp.350. ISBN 90-277-1814-8. £31.75, \$49.50.

The Universe of Galaxies: Readings from Scientific American. PAUL W. HODGE (ed.). W.H. Freeman: 1984. Pp.113. Pbk ISBN 0-7167-1676-3. Pbk £10.95.

Physics

Acoustooptical Phenomena in Liquid Crystals. By O.A. KAPUSTINA. Gordon and Breach: 1984. Pp.164. ISBN 0-677-06605-8. \$65.

Analytical Laser Spectroscopy. S. MARTELLUCCI and A.N. CHESTER (eds). Plenum: 1985. Pp.290. ISBN 0-306-41897-5. Np.

Antennas and Waveguides for Nonsinusoidal Waves. By HENNING F. HARMUTH. Academic: 1984. Pp.276. ISBN 0-12-014577-4. \$60, £45.

Applied N=1 Supergravity. By P. NATH, R. ARNOWITT and A.H. CHAMSEDDINE. Wiley: 1984. Pp.116. Pbk ISBN 9971-966-49-2. Pbk £8.80.

Astrophysics Today: Readings from Physics Today. A.G.W. CAMERON (ed.). American Institute of Physics: 1984. Pp.337. Pbk ISBN 0-88318-446-X. \$25.

Auger Electron Spectroscopy. By MICHAEL THOMPSON et al. Wiley: 1985. Pp.394. ISBN 0-471-04377-X. £95.

Basic Concepts in Relativity and Early Quantum Theory, 2nd Edn. By ROBERT RESNICK and DAVID HALLIDAY. Wiley: 1985. Pp.341. ISBN 0-471-88858-3. £19.95.

Chaos in Nonlinear Dynamical Systems. JAGDISH CHANDRA (ed.). SIAM/Wiley: 1984. Pp.191. ISBN 0-89871-052-9. £23.80.

Condensed Matter Research Using Neutrons: Today and Tomorrow. STEPHEN W. LOVESEY and REINHARD SCHERM (eds). Plenum: 1984. Pp.329. ISBN 0-306-41821-5. Np.

Electrodynamics of Continuous Media, 2nd Edn. By L.D. LANDAU, E.M. LIFSHITZ and L.P. PITAEVSKII. Pergamon: 1984. Pp.460. Pbk ISBN 0-08-030275-0. Pbk £15.50, \$25.

Ferroelectrics. R. BLINC et al. (eds). Gordon and Breach: 1984. Pp.499. Pbk ISBN 0-667-16595-1. Pbk \$130.

Fifth Workshop on Grand Unification. KYUNG-SIK KANG et al. (eds). World Scientific: 1984. Pp.538. ISBN 9971-966-58-1. £51.75.

Fundamentals of Microoptics: Distributed-Index, Microlens, and Stacked Planar Optics. K. IGA et al. (eds). Academic: 1985. Pp.218. ISBN 0-12-370360-3. \$38, £29.50.

Heat Transfer, Vols 1 & 2. INSTITUTION OF CHEMICAL ENGINEERS. Pergamon: 1984. Vol. 1 ISBN 0-85295-174-4. Vol. 2 ISBN 0-85295-175-2. Pp.1322. £70, \$112.

Infrared and Millimeter Waves, Vol. 12. Electromagnetic Waves in Matter. KENNETH J. BUTTON (ed.). Academic: 1984. Pp.323. ISBN 0-12-147712-6. \$69, £52.

Inverse Problems of Acoustic and Elastic Waves. FADIL SANTOSA et al. (eds). SIAM/Wiley: 1984. Pp.365. ISBN 0-89871-050-2. £36.65.

James E. Keeler: Pioneer American Astrophysicist and the Early Development of American Astrophysics. By DONALD E. OSTERBROCK. Cambridge University Press: 1984. Pp.411. ISBN 0-521-26582-7. £35.

Leptons, Hadrons and Nuclei. By FLORIAN SCHECK. North-Holland: 1983. Pp.388. ISBN 0-444-86719-8. \$61.50.

Line Coincidence Tables for Inductively Coupled Plasma Atomic Emission Spectrometry, Vols 1 & 2, 2nd Edn. By P.W.J.M. BOUMANS. Pergamon: 1984. ISBN 0-31404-X. £135, \$215.

Magnetic Excitations and Fluctuations. S.W. LOVESEY et al. (eds). Springer-Verlag: 1984. Pp.227. ISBN 3-540-13789-0/0-387-13789-0. DM80, \$28.10.

Multiphoton Processes. P. LAMBROPOULOS and S.J. SMITH (eds). Springer: 1984. Pp.201. ISBN 0-387-15068-4. Np.

Nonequilibrium Cooperative Phenomena in Physics and Related Fields. MANUEL G. VELARDE (ed.). Plenum: 1984. Pp.569. ISBN 0-306-41833-9. Np.

Particle Physics: The Quest for the Substance of Substance. By L.B. OKUN. Gordon & Breach: 1985. Pp.223. Hbk ISBN 3-7186-0228-8; pbk ISBN 3-7186-0229-6. Hbk \$44; pbk \$18.

Principles of Plasma Electrodynamics. By A.F. ALEXANDROV, L.S. BOGDANKEVICH and A.A. RUKHADZE. Springer-Verlag: 1984. Pp.488. ISBN 3-540-12613-9/0-387-12613-9. DM135.

Quantum Field Theory. By LEWIS H. RYDER. Cambridge University Press: 1985. Pp.443. ISBN 0-521-23764-5. £40, \$74.50.

Seismic Migration: Imaging of Acoustic Energy by Wave Field Extrapolation. By A.J. BERKHOUT. Elsevier: 1984. Pp.274. ISBN 0-444-42431-8. Np.

Simulation for Nuclear Reactor Technology. By D.G. WALTON. Cambridge University Press: 1985. Pp.444. ISBN 0-521-26785-4. £35, \$69.50.

Solitons in Nuclear and Elementary Particle Physics. A. CHODOS, E. HADJIMICHAEL and C. TZE (eds). World Scientific: 1984. Pp.313. ISBN 9971-966-89-1. £39.10.

Space, Time and Causality: An Essay in Natural Philosophy. By J.R. LUCAS. Clarendon: 1984. Pp.206. ISBN 0-19-875057-9. £19.50.

Surface Enhanced Raman Vibrational Studies at Solid/Gas Interfaces. By I. POCKRAND. Springer-Verlag: 1984. Pp.187. ISBN 3-540-13671-1/0-387-13416-6. DM75, \$26.40.

Theory of Slow Atomic Collisions. By E.E. NIKITIN and S. Ya. UMANSKII. Springer-Verlag: 1984. Pp.432. ISBN 3-540-12414-4/0-387-12414-4. DM 140.

Thin Film and Depth Profile Analysis. H. OECHSNER (ed.). Springer-Verlag: 1984. Pp.205. ISBN 3-540-13320-8/0-387-13320-8. DM72.

Time-Correlated Single Photon Counting. By DESMOND V. O'CONNOR and DAVID PHILLIPS. Academic: 1984. Pp.288. ISBN 0-12-524140-2. \$45, £31.

Treatise on Heavy-Ion Science, Vols 2 & 3. D. ALLAN BROMLEY (eds). Plenum: 1984. Vol. 2: pp.734, ISBN 0-306-41572-0. Vol. 3: pp.589, ISBN 0-306-41573-9. Np.

Chemistry

Advances in Heterocyclic Chemistry. ALAN R. KATRITZKY FRS (ed.) Academic: 1984. Pp.368. ISBN 0-12-020637-4. \$85, £65.

Advances in Organometallic Chemistry, Vol. 23. F.G.A. STONE and R. WEST (eds). Academic: 1984. Pp.319. ISBN 0-12-031123-2. \$65, £49.

Advances in Physical Organic Chemistry, Vol. 20. V. GOLD and D. BETHELL (eds). Academic: 1984. Pp.252. ISBN 0-12-033520-4. \$50, £35.

Analytical Measurement and Information: Advances in the Information Theoretic Approach to Chemical Analyses. By K. ECKSCHLAGER and V. ŠTĚPANEK. Wiley: 1985. Pp.140. ISBN 0-86380-021-1. £25.

The Central Science: Essays on the Uses of Chemistry. G.B. KAUFFMAN and H.H. SZMANT (eds). Texas Christian University Press: 1984. Pp.181. ISBN 0-912646-84-5. \$15.

Chemical Mössbauer Spectroscopy. R.H. HERBER (ed.). Plenum: 1984. Pp.378. ISBN 0-306-41885-1. Np.

Chemistry and Physics of Carbon. PETER A. THROWER (ed.). Dekker: 1984. Pp.360. ISBN 0-8247-7245-8. \$79.50.

Dictionary of Electrochemistry, 2nd Edn. By D.B. HIBBERT and A.M. JAMES. Macmillan, London: 1984. Pp.308. ISBN 0-333-34983-0. £20.

Encyclopedia of Electrochemistry of the Elements, Vol. 15. ALLEN J. BARD and HENNING LUND (eds). Dekker: 1984. Pp.324. ISBN 0-8247-2515-8. \$129.50, SwFr.324.

Food: The Chemistry of its Components. By T.P. COULTATE. Royal Society of Chemistry: 1984. Pp.197. Pbk ISBN 0-85186-483-X. Pbk £5.95.

Handbook of Chemical Equilibria in Analytical Chemistry. By S. KOTRLÝ and L. ŠŮCHA. Ellis Horwood: 1985. Pp.414. ISBN 0-85312-107-9. £45.

Infrared Spectroscopy in Chemical Analysis: An Interactive Software Package. By D. KEALEY. Wiley: 1985. Np.

Metallo-Organic Chemistry. By ANTHONY J. PEARSON. Wiley: 1985. Pp.398. Pbk ISBN 0-471-90446-5. Pbk £9.95.

Mössbauer Spectroscopy Applied to Inorganic Chemistry, Vol. 1. GARY J. LONG (ed.). Plenum: 1984. Pp.667. ISBN 0-306-41647-6. Np.

Nuclear Analytical Chemistry. By DAG BRUNE, BENGT FORKMAN and BERTIL PERSSON. Charwell-Bratt, Old Orchard, Bickley Road, Bickley, Kent: 1984. Pp.557. ISBN 0-86238-0472-2. Np.

The Organic Chemistry of Drug Synthesis, Vol. 3. By D. LEDNICER and L.A. MITSCHER. Wiley: 1984. Pp.284. ISBN 0-471-09250-9. £37.95.

Organic Reaction Mechanisms, 1983. A.C. KNIPE and W.E. WATTS (eds). Wiley: 1985. Pp.589. ISBN 0-471-90503-8. £78.50.

Polymer Processing and Properties. GIANNI ASTARITA and LUIGI NICOLAIS (eds). Plenum: 1984. Pp.450. ISBN 0-306-41728-6. Np.

Practical Aspects of Gas Chromatography/Mass Spectrometry. By GORDON M. MESSAGE. Wiley: 1984. Pp.351. ISBN 0-471-06277-4. £69.35.

Statistics for Analytical Chemistry. By J.C. MILLER and J.N. MILLER. Ellis Horwood: 1984. Pp.202. ISBN 0-85312-662-3. £18.50.

A Symmetric Synthesis: Vol. 3 Stereodifferentiating Addition Reactions. JAMES D. MORRISON (ed.). Academic: 1984. Pp.578. ISBN 0-12-507703-3. \$69.50.

Techniques and Application of Thin Layer Chromatography. JOSEPH C. TOUCHSTONE and JOSEPH SHERMA (eds). Wiley: 1985. Pp.395. ISBN 0-471-88017-5. £88.65.

Techniques in Organic Reaction Kinetics. By PETR ZUMAN and RAMESH PATEL. Wiley: 1984. Pp.340. ISBN 0-471-03556-4. £57.25.

Computer Science

Computational Functional Analysis. By R.E. MOORE. Ellis Horwood: 1985. Pp.155. ISBN 0-85312-807-3. £15.

Computer Culture: The Scientific, Intellectual, and Social Impact of the Computer. HEINZ R. PAGELS (ed.). New York Academy of Sciences: 1984. Pp.288. Pbk ISBN 0-89766-245-8. Pbk \$66.

Distributed Computing Systems Programme. D. A. DUCE (ed.). *Peter Peregrinus*: 1984. Pp.305. ISBN 0-86341-023-5. £15, \$38.

Effective Fortran 77. By MICHAEL METCALF. *Oxford University Press*: 1985. Pp.231. Pbk ISBN 0-19-853709-3. Pbk £9.95.

Fundamentals of Human-Computer Interaction. ANDREW MONK (ed.). *Academic*: 1984. Pp.293. ISBN 0-12-504580-8. \$26.50, £18.50.

A Glossary of Computing Terms, 4th Edn. THE BRITISH COMPUTER SOCIETY. *Cambridge University Press*: 1984. Pp.58. Pbk ISBN 0-521-31778-9. Pbk £1.35.

Image Processing System Architectures. JOSEF KITTLER and MICHAEL J.B. DUFF (eds). *Research Studies Press*: 1985. Pp.166. ISBN 0-86380-023-8. £19.75.

The Microcomputer Users Handbook 1985. By DENNIS LONGLEY and MICHAEL SHAIN. *Macmillan*, London: 1985. Pp.405. Pbk ISBN 0-333-36866-5. Pbk £4.99.

New Computer Architecture. J. TIBERGHEN (ed.). *Academic*: 1984. Pp.289. ISBN 0-12-690980-6. \$28.50, £19.50.

Silicon Shock: The Menace of the Computer Invasion. By GEOFF SIMONS. *Basil Blackwell*: 1985. Pp.191. ISBN 0-631-13835-8. £9.50.

Technology

Advances in Electrochemistry and Electrochemical Engineering, Vol. 13. By HEINZ GERISCHER and CHARLES W. TOBIAS. *Wiley*: 1984. Pp.360. ISBN 0-471-87881-2. £52.

Creep of Crystals: High-Temperature Deformation Processes in Metals, Ceramics and Minerals. By JEAN-PAUL POIRIER. *Cambridge University Press*: 1985. Pp.260. Pbk ISBN 0-521-27851-1. Pbk £10.95.

Electronic Synthesis of Speech. By R. LINGGARD. *Cambridge University Press*: 1985. Pp.149. ISBN 0-521-24469-2. £15, \$29.95.

Engineering Flow and Heat Exchange. By OCTAVE LEVENSPIEL. *Plenum*: 1984. Pp.366. ISBN 0-306-41599-2. \$34.50.

The Information Technology Revolution. TOM FORESTER (ed.). *Basil Blackwell*: 1985. Pp.674. Pbk ISBN 0-631-13438-7. Pbk £9.95.

Earth Sciences

Archaeological Geology. GEORGE RAPP Jr and JOHN A. GIFFORD (eds) *Yale University Press*: 1985. Pp.435. ISBN 0-300-03142-4. £35.

Coal Combustion and Gasification. By L. DOUGLAS SMOOT and PHILIP J. SMITH. *Plenum*: 1985. Pp.443. ISBN 0-306-41750-2. Np.

Coastline Changes: A Global Review. By ERIC C.F. BIRD. *Wiley*: 1985. Pp.219. ISBN 0-471-90646-8. £22.50.

Elements of Petroleum Geology. By RICHARD C. SELLEY. W.H. Freeman: 1985. Pp.449. ISBN 0-7167-1630-5. \$44.95.

Feldspars and Feldspathoids: Structures, Properties and Occurrences. W. L. BROWN (ed.) *Reidel*: 1984. Pp.541. ISBN 90-277-1826-1. \$74, £49.50.

Hemispherical Projection Methods in Rock Mechanics. By S.D. PRIEST. *George Allen & Unwin*: 1985. Pp.124. Pbk ISBN 0-04-622007-0. Pbk £8.95.

Phanerozoic Earth History of Australia. By J.J. VEEVERS. *Oxford University Press*: 1985. Pp.418. ISBN 0-19-854459-6. £55.

Rocks and Minerals. By PAT BELL and DAVID WRIGHT. *Hamlyn*: 1985. Pp.192. Pbk ISBN 0-600-30556-2. Pbk £3.99.

Sedimentation on the Continental Shelf with Special Reference to the East China Sea. ACTA OCEANOLOGICA SINICA (ed.). *Springer-Verlag*: 1983. Pp.952. ISBN 3-540-12648-1. DM 475.

Environmental Sciences

Air Pollution Modelling and its Applications, Vol. 4. NATO Challenges of Modern Society, Vol. 7. C. DE WISPELAERE (ed.). *Plenum*: 1985. Pp.818. ISBN 0-306-41908-4. Np.

Causes and Effects of Changes in Stratospheric Ozone: Update 1983. COMMITTEE ON CAUSES AND EFFECTS OF CHANGES IN STRATOSPHERIC OZONE. *National Academy*: 1984. Pp.254. ISBN 0-309-03443-4. Np.

Environmental Evaluation: Perception and Public Policy. By ERVIN H. ZUBE. *Cambridge University Press*: 1985. Pp.148. ISBN 0-521-31972-2. Pbk £5.95.

Environmental Science: A Framework for Decision Making. By DANIEL D. CHIRAS. *Benjamin/Cummings*: 1985. Pp.655. ISBN 0-8053-2255-8. \$33.95.

Social and Economic Aspects of Radioactive Waste Disposal: Considerations for Institutional Management. PANEL ON SOCIAL AND ECONOMIC ASPECTS OF RADIOACTIVE WASTE MANAGEMENT. *National Academy*: 1984. Pp.175. ISBN 0-309-03444-2. Np.

Technical Report Series, Vol. 718: Environmental Pollution Control in Relation to Development. WORLD HEALTH ORGANISATION EXPERT COMMITTEE. *World Health Organisation*: 1985. Pp.63. ISBN 0-924-1207187-3. SwFr.6.

Biological Sciences

Acquired Immunodeficiency Syndrome (AIDS): Advances in Host Defense Mechanisms, Vol. 5. JOHN I. GALLING and ANTHONY S. FAUCI (eds). *Raven*: 1985. Pp.178. ISBN 0-88167-077-4. \$42.

Adaptation to Altitude-Hypoxia in Vertebrates: Zoophysiology, Vol. 16. By PIERRE BOUVEROT. *Springer-Verlag*: 1985. Pp.176. ISBN 3-540-13602-9/0-387-13602-9. DM94.

Advances in Anatomy Embryology and Cell Biology, Vol. 86. By JEFFREY A. WINER. *Springer-Verlag*: 1984. Pp.97. ISBN 3-540-13524-6. DM58.

Advances in Anatomy Embryology and Cell Biology, Vol. 92. Angiographic Anatomy of the Anterior Inferior Cerebellar Artery. By JAN J. HEIMANS, JAAP VALK and ANTHONY H.M. LOHMAN. *Springer-Verlag*: 1985. Pp.91. ISBN 3-540-13768-8/0-387-13768-8. DM58.

Advances in Cardiology, Vol. 32. Assessment of Ventricular Function. D.H. SPODICK (ed.). *Karger*: 1985. Pp.153. ISBN 3-8055-3993-2. SwFr126, DM151, \$53.75.

Advances in Human Genetics. H. HARRIS and K. HIRSCHHORN (eds). *Plenum*: 1985. Pp.399. ISBN 0-306-41752-9. Np.

Advances in Microbial Ecology. K.C. MARSHALL (ed.). *Plenum*: 1985. Pp.307. ISBN 0-306-41877-0. Np.

Advances in Myocardiology 6. N.S. DHALLA and D.J. HEARSE (eds). *Plenum*: 1985. Pp.669. ISBN 0-306-41759-6. Np.

Algal Cell Biology. By H.D. KUMAR. *East-West Press*: 1985. Pp.201. Pbk Rs 41.

Animal Behaviour. By DAVID MCFARLAND. *Pitman*: 1985. Pp.576. ISBN 9-780273-021032. £11.95.

Animal Intelligence. L. WEISKRANTZ (ed.). *Oxford University Press*: 1985. Pp.223. ISBN 0-19-852124-3. £32.50.

Auditory Development in Infancy. SANDRA E. TREHUB and BRUCE SCHNEIDER (eds). *Plenum*: 1985. Pp.243. ISBN 0-306-41757-X. Np.

Autonomic Functions in Human Physiology. By G. THEWS and P. VAUPEL. *Springer-Verlag*: 1985. Pp.384. ISBN 3-540-13217-1/0-387-13217-1. DM39.80, \$14.

Structural and Resonance Techniques in Biological Research. DENIS L. ROUSEAU (ed.). *Academic*: 1984. Pp.476. ISBN 0-12-599320-X. \$69, £52.

Surgical Management of the Burn Wound. By DAVID M. HEIMBACK and LOREN H. ENGRAV. *Raven*: 1984. Pp.161. ISBN 0-89004-914-9. \$93.50.

Synthetic Pyrethroids and Other Pesticides. G. ZWEIG and J. SHERMA (eds). *Academic*: 1984. Pp.312. ISBN 0-12-784313-2. \$65, £49.

Transnasal Systemic Medications: Fundamentals, Developmental Concepts and Biomedical Assessments. YIE W. CHIEN (ed.) *Elsevier*: 1985. Pp.226. ISBN 0-444-42460-1. Dfl145, \$53.75.

Two-Dimensional Gel Electrophoresis of Proteins: Methods and Applications. JULIO E. CELIS and RODRIGO BRAVO (eds). *Academic*: 1984. Pp.487. ISBN 0-12-164720-X. \$69.50.

Understanding Dental Caries: Etiology and Mechanisms. By GORDON NIKIFORUK. *Karger*: 1985. Pp.306. ISBN 3-8055-3864-2. DM69, \$34.75.

Wildlife Pest Control Around Gardens and Homes. By TERRELL P. SALMON and ROBERT E. LICKLITER. *University of California Division of Agriculture and Natural Resources*: 1984. Pp.90. Pbk ISBN 0-931876-66-4. Np.

Archaeology

Excavations at Grimes Graves Norfolk 1972 - 1976. By JULIE CLUTTON-BROCK. *British Museum Publications*: 1985. Pp.46. Pbk ISBN 0-7141-1374-3. Pbk £10.

The Greek Trireme of the 5th Century B.C. JOHN COATES and SEAN MCGRAIL (eds). *National Maritime Museum, Greenwich*: 1985. Pp.140. Pbk ISBN 0-905555-95-3. Pbk £8.

Paleopathology at the Origins of Agriculture. MARK NATHAN COHEN and GEORGE J. ARMELAGOS (eds). *Academic*: 1984. Pp.615. ISBN 0-12-179080-0. \$59, £41.50.

The Palynology of Archaeological Sites. By GEOFFREY W. DIMBLEBY. *Academic*: 1985. Pp.176. ISBN 0-12-216480-6. \$45, £46.

General

Handbook of Discourse Analysis, Vol.1. Disciplines of Discourse. TEUN. A. VAN DIJK (ed.). *Academic*: 1985. Pp.302. ISBN 0-12-712001-7. \$59.50, £45.50.

High Altitude Deterioration. J. RIVOLIER *et al.* (eds). *Karger*: 1985. Pp.228. ISBN 3-8055-3972-X. DM 174, \$61.75.

Ideas, Qualities and Corpuscles: Locke and Boyle on the External World. By PETER ALEXANDER. *Cambridge University Press*: 1985. Pp.336. ISBN 0-521-26707-2. £27.50.

Image Understanding 1984. SHIMON and WHITMAN RICHARDS (eds). *Ablex Publishing*, 355 Chestnut Street, Norwood, New Jersey: 1985. Pp.268. ISBN 0-89391-254-9. \$35.

Modern Trends in Hypnosis. DAVID WAXMAN *et al.* (eds). *Plenum*: 1985. Pp.429. ISBN 0-306-41905-X. Np.

Multivariable Control: New Concepts and Tools. SPYROS G. TZAFESTAS (ed.). *Reidel*: 1984. Pp.502. ISBN 90-277-1829-6. Dfl. 1230, \$49.50, £33.25.

Nuclear Fact Book, 2nd Edn. A.M. PLATT, J.V. ROBINSON and O.F. HILL (eds). *Harwood Academic*: 1985. Pp.176. ISBN 3-7186-0273-3. \$33.

On the Frontier: Flight Research at Dryden, 1946-1981. By RICHARD P. HALLION. *NASA*: 1984. Pp.385. Np.

Proceedings of the Royal Institution of Great Britain, Vol. 56. NINA HALL and IVOR WILLIAMS (eds). *Science Reviews*: 1984. Pp.288. ISBN 0-905927-81-8. Np.

Public Health Papers, Vol. 81: Community Response to Alcohol-related Problems: Review of an International Study. By E.B. RITSON. *World Health Organisation*: 1985. Pp.58. ISBN 92-4-130081-7. Np.

The Regional Economic Impact of Technological Change. A.T. THWAITES and R.P. OAKLEY (eds). *Frances Pinter*: 1985. Pp. 249. ISBN 0-86187-373-4. £17.50.

Research: How to Plan, Speak and Write About it. CLIFFORD HAWKINS and MARCO SORGI (eds). *Springer-Verlag*: 1985. Pp.184. ISBN 0-387-13992-3/ ISBN 3-540-13992-3. DM 32.

Revelations: Glimpses of Reality. RONALD S. LELLO (ed.). *Shepherd-Walwyn*: 1985. Pp.136. ISBN 0-85683-079-8. £8.50.

Safeguarding the Atom: A Critical Appraisal. JOZEF GOLDBLAT (ed.). *SIPRI*: 1985. Pp.243. ISBN 0-85066-306-7. Np.

Science and Technology Policies in Finland and Hungary: A Comparative Study. K.O. DONNER and L. PAL (eds). *Akadémi Kiadó, Budapest*: 1985. Pp.371. ISBN 963-05-3977-2. £19.75.

United States Arctic Interests: The 1980s and 1990s. WILLIAM E. WESTERMEYER and KURT M. SHUSTERICH (eds). *Springer-Verlag*: 1984. Pp.369. ISBN 0-387-96009-0. DM 114.